DOCUMENT RESUME

ED 231 390 IR 050 316

AUTHOR	Markey, Karen
TITLE	The Process of Subject Searching in the Library Catalog: Subject Access Research Project. Final Report.
INSTITUTION	OCLC Online Computer Library Center, Inc., Dublin, Ohio.
REPORT NO	OCLC/OPR/RR-83-1
PUB DATE	4 Feb 83
NOTE	159p.
PUB TYPE	Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	*Design Requirements; Information Retrieval; Interviews; Library Automation; *Library Catalogs; Library Research; Man Machine Systems; *Models; *Online Systems; Questionnaires; *Search Strategies; Use Studies
IDENTIFIERS	*Library Users; *Online Catalogs

ABSTRACT

The Subject Access Research Project was initiated by OCLC in August 1979 to study the process of subject searching in the library catalog, with the objective of identifying features for an automated subject retrieval system which would support the present search tactics employed by library users performing manual subject searches at the library catalog. Subject searchers' spoken thoughts (i.e., protocols) were recorded as they performed their subject searches at varied libraries. An analysis of protocols resulted in construction of a subject searching process model. Features suggested by the model are recommended for implementation in an online public access catalog (OPAC), which has the potential to support new subject-searching features. An OPAC in which the searcher's model of the information retrieval process of traditional catalog searching is implemented may have the same potential as the library catalog, but inclusion of a user-oriented interface is necessary to make it as powerful. An OPAC with such an interface and enhanced with features not previously available to library catalog searchers would actually be more powerful than the traditional catalog. This report includes 39 figures, 27 tables and a 29-item reference list as well as appendices including data collection instruments. (Author/LMM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED231390

Report Number: OCLC/OPR/RR-83/1
1983 February 4

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Research Report

on

THE PROCESS OF SUBJECT SEARCHING IN THE LIBRARY CATALOG: FINAL REPORT OF THE SUBJECT ACCESS RESEARCH PROJECT

by

Karen Markey

OCLC Online Computer Library Center, Inc.
Office of Research
6565 Frantz Road
Dublin, OH 43017

IR050314

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Lois Yoakam

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

All rights reserved; however, individuals may quote from or reproduce portions of the text for noncommercial, educational, or private study or research. Full credit must be given to both the author(s) and the Office of Research, OCLC, Dublin, Ohio.

The Research Report Series is OCLC's formal dissemination vehicle through which OCLC research project results can be made public.

OCLC Research Reports are published and distributed by OCLC until the report is available from ERIC (usually six to nine months after publication). Up to three copies of reports not yet bearing an ERIC number are provided at no charge to interested persons and institutions. Requests for these reports should be directed to OCLC, Documentation Department, 6565 Frantz Road, Dublin, OH 43017.

Research Report Series

Subject Heading Patterns in OCLC Monographic Records
by E. O'Neill and R. Aluri
Report Number: OCLC/RDD/RR-79/1; ERIC ED 183167

An Overview of a Proposed Monitoring Facility for the Large-scale, Network-based OCLC Online System
by W. Dominick, D. Penniman, and J. Rush
Report Number: OCLC/OPR/RR-80/1; ERIC ED 186042

Analytical Review of Catalog Use Studies
by K. Markey
Report Number: OCLC/OPR/RR-80/2; ERIC ED 186041

A Method for Correcting Typographical Errors in Subject Headings in OCLC Records
by E. O'Neill and R. Aluri
Report Number: OCLC/OPR/RR-80/3; ERIC ED 215680

Field, Subfield, and Indicator Statistics in OCLC Bibliographic Records
by T. B. Hickey
Report Number: OCLC/OPR/RR-81/1; ERIC ED 202484

Development of a Probabilistic Author Search and Matching Technique for Retrieval and Creation of Bibliographic Records
by T. B. Hickey
Report Number: OCLC/OPR/RR-81/2; ERIC ED 206286

Survey of Online Searching Instruction in Schools of Library and Information Science
by C. H. Fenichel and S. P. Harter
Report Number: OCLC/DD/RR-81/3; ERIC ED 210038

The American Public Library
by D. Sager
Report Number: OCLC/OPR/RR-82/1; ERIC ED 221222

ACKNOWLEDGMENTS

Many individuals have made substantial contributions to this Project, the first of course being Frederick G. Kilgour who, when Director of OCLC, realized the impending importance of subject access and, in summer 1979, charged the Research Department with the Subject Access Research Project. Dr. W. David Penniman identified the scope and methodology of the Project, and data collection was organized and supervised by Dr. Neal K. Kaske. Participating libraries are thanked for their support in the data collection phase. Nancy Sanders and Patricia Hartsell assisted in collecting data. Patricia Hartsel drew up and applied the coding scheme and helped perform the data analysis; her assistance is especially acknowledged. The extensive editorial assistance offered by Ginni Voedisch was invaluable in preparing the manuscript; the editorial advice of Sig F. Dierk and Lois L. Yoakam was also helpful. The typing assistance of Cindy Heffelfinger, who prepared the rough drafts of the text and the final copies of the protocols and tables, was particularly helpful. Numerous colleagues provided critical enlightenment and were themselves sources of information, particularly Daniel Ferrer, Charles R. Hildreth, Neal K. Kaske, William H. Mischo, and W. David Penniman. Professor Pauline A. Cochrane provided support in both the data analysis and interpretation phases of the Project and her review and criticism of early drafts of the report was especially helpful.

NOTES ABOUT THE AUTHOR

Karen Markey is a Research Scientist in OCLC's Office of Research. She has conducted research in online database coverage, subject access to augmented MARC records, and the description of subject matter in works of art. She is currently Project Manager of the evaluation phase of the Online Public Access Catalog Evaluation Project sponsored by the Council on Library Resources. She received her B.A. from The Johns Hopkins University and her M.L.S. and Ph.D. from the School of Information Studies at Syracuse University.

ABSTRACT

This report describes the results of the Subject Access Research Project which was initiated by OCLC in August 1979 to study the process of subject searching in the library catalog. The major objective of the project was to identify features for an automated subject retrieval system which support the present search tactics employed by library users performing manual subject searches at the library catalog. Subject searchers' spoken thoughts (i.e., protocols) were recorded as they performed their subject searches. An analysis of protocols resulted in the construction of a model of the process of subject searching. Features suggested by the model are recommended for implementation in an online public access catalog (OPAC) to support the search tactics presently employed by searchers of the traditional library catalog. In addition, thanks to technology, the OPAC has the potential to support features never before possible in subject searching. The user interface--a new dimension in catalog use available in the OPAC--is the crucial communications component of the OPAC through which the searcher unleashes the system's power. An OPAC in which the searcher's model of the information retrieval process of traditional catalog searching is implemented may have the same potential as the library catalog, but inclusion of a user-oriented interface is necessary to make that OPAC as powerful as the library catalog. An OPAC enhanced with features not previously available to library catalog searchers may have more potential than the traditional catalog, but only the inclusion of a user-oriented interface actually makes that OPAC more powerful than the library catalog.

CONTENTS

ABSTRACT v

LIST OF FIGURES ix

LIST OF TABLES xii

EXECUTIVE SUMMARY 1

I. PROJECT OBJECTIVES: A MODEL OF SUBJECT SEARCHING AT THE LIBRARY CATALOG 14

II. METHODOLOGY 16

A. Protocol Analysis for Data Collection 16
B. Coding Scheme for Analysis of Protocols 19

III. DATA ANALYSIS 25

A. Participating Libraries 25
B. Subject Searchers at Participating Libraries 25
C. Frequency of Constituent Elements in Subject Searches 28
D. Access Points in Subject Searches 42
E. Scans in Subject Searches 60
F. Refinements in Subject Searches 62
G. Completing the Subject Search: Aids, Flags, Records, and Exits 75
H. Conclusion 79

IV. DISCUSSION: AD HOC PROCESS MODELS OF SUBJECT SEARCHING AND IMPLICATIONS FOR POWER 82

A. Introduction: Ad Hoc Process Model of Subject Searching 82
B. Bookshelf Browsing Subject Searches 84
   1. Description 84
   2. Required Features 84
   3. Beyond the Library Catalog: Bookshelf Browsing Features for OPACs 84
C. Subject Searches with a priori Refinements 87
   1. Description 87
   2. Required Features 91
   3. Beyond the Library Catalog: a priori Refinements in Subject Searches 91
D. Subject Searches with Interactive Refinements 92
   1. Description 92
   2. Required Features 95
   3. Beyond the Library Catalog: Interactive Refinements in Subject Searches 99
E. Subject Searches, Multiple Access Points, and Assistance 100
   1. Description 100
   2. Required Features 100
   3. Beyond the Library Catalog: Multiple Access Points and Assistance in Subject Searches 105

CONTENTS

F. The User-Oriented Interface in Online Public Access Catalogs 111

REFERENCES 116

APPENDIX A. Agenda: Training for Protocol Analysis Interviews 119

APPENDIX B. Protocol Analysis Interview Questionnaire 122

APPENDIXES C1-C7. Expressed Topics and Access Points 124

INDEX 135

LIST OF FIGURES

1 Flowchart of the Major Tasks of Study Areas 3-5 of the Subject Access Research Project 17

2 Coded Protocol of a Subject Search for Financial Statement Analysis 24

3 Coded Protocol of a Subject Search for Christmas in Germany 33

4 Frequency of Constituent Codes and Number of Searches Assigned Each Frequency 34

5 Coded Protocol of a Subject Search for Information Theory Assigned Four Constituent Elements 35

6 Subject Guide Printed Display of Subject Headings, Class Numbers, and Number of Books per Class Number 35

7 Extracts of the Dewey Decimal Classification Schedules where Books Are Assigned the Subject Heading "Information Theory" 37

8 Coded Protocol of a Bookshelf Browsing Search for Real Estate 41

9 Partially Coded Protocol of a Subject Search for Canada's Anthem (Initial Access Point Is a Partial Representation of Expressed Topic--1A) 51

10 Partially Coded Protocol of a Subject Search for How Good a Breakfast is (Initial Access Point Is Different from Expressed Topic--1C) 54

11 Partially Coded Protocol of a subject Search for Motorcycles (Initial Access Point Is the Same as Expressed Topic--1B) 57

12 Partially Coded Protocol of a Subject Search for Mondrian (Access Point Is the Same as Expressed Topic--1B) 58

13 Example of an Intermediate Display of a Retrieved Item in an OPAC Subject Search 63

14 Protocol of a Subject Search for Trance Dances of the Holy Ghost People Coded with Scans and Access Points 64

15 Coded Extract of a Protocol of a Subject Search for Airships with a priori Refinements 68

16 Coded Protocol of a Subject Search for Prenatal Care with a priori and Interactive Refinements 70

17 Coded Protocol of a Subject Search for Small Business with a priori and Interactive Refinements 74

18 Ad Hoc Process Model--Composite Model 83

19 Ad Hoc Process Model of a Bookshelf Browsing Subject Search 85

LIST OF FIGURES

20 Alphabetical Subject Heading Display with Class Numbers for OPAC Bookshelf Browsers 86

21 Suggested Online Display of the Subject Guide to the Collection of the Public Library of Columbus and Franklin County 86

22 Coded Protocol of a Subject Search for Audi Repair Manual with a priori Refinements 88

23 Ad Hoc Process Model of a Subject Search for Audi Repair Manual with a priori Refinements 89

24 OPAC Search for Automobiles, Foreign in Maggie's Place (the OPAC of Pikes Peak Regional Library District, Colorado) 90

25 Coded Protocol of a Subject Search for Multinational Corporations with Interactive Refinements 93

26 Ad Hoc Process Model of a Subject Search for Multinational Corporations with Interactive Refinements 94

27 OPAC Search for Multinational Corporations in Mankato State's MSUS/PALS Featuring Explicit Boolean Operations 96

28 OPAC Search for Multinational Corporations in Syracuse's SULIRS Featuring Implicit Boolean Operations 98

29 Coded Protocol of a Subject Search for FCC with Multiple Access Points and No Records (i.e., Failure) 101

30 Ad Hoc Process Model of a Subject Search for FCC with Multiple Access Points and No Records (i.e., Failure) 102

31 Coded Protocol of a Subject Search (with Librarian Assistance) for Life after Death 103

32 Ad Hoc Process Model of a Subject Search (with Librarian Assistance) for Life after Death 104

33 Related Term Display with Class Areas Constructed from the Printed Library of Congress Subject Headings 106

34 Keyword-in-Context Display of Subject Heading (Islam) from the University of California's MELVYL 107

35 Keyword-in-Context Display of Subject Heading (Islam) with Geographical Subdivisions 108

36 Online Vocabulary Aid for Suggesting to OPAC Searchers Subject Headings and Class Number Areas Common to Displayed Items 110

LIST OF FIGURES

37 The Dewey Decimal Classification Scheme in the Online Public Access Catalog 110

38 Guidelines for OPAC Evaluation and Improvement 113

39 Necessary Components of OPACs with as Much or More Power than Traditional Library Catalogs 115

LIST OF TABLES

1	Search Activity Codes and Their Constituent Elements	21
2	Participating Libraries	26
3	Age of Subject Searchers at Participating Libraries	27
4	Highest Grade Completed by Subject Searchers at Participating Libraries	27
5	Frequency of Library Visits by Subject Searchers at Participating Libraries	27
6	Percentage of Subject Searchers Reporting Some Training in Searching a Library Catalog	30
7	Purpose of Subject Searchers at Participating Libraries	30
8	Frequency of Search Activity Codes and Their Constituent Elements	31
9	Chi-Square Test of the Relationship between Number of Access Points and Constituent Elements	39
10	Chi-Square Test of the Relationship between Number of Constituent Elements and Type of Library	40
11	Frequency of Access Points in Subject Searches	44
12	Number of Searchers and Number of Access Points	44
13	Perseverance and Type of Initial Access Point	45
14	Chi-Square Test of the Relationship between Type and Number of Initial Access Points	46
15	Percentage and Type of Access Points Occurring as Initial, Middle, and Last Access Points	47
16	Chi-Square Test of the Relationship between Number of Access Points and Request for Librarian Assistance	47
17	Number of Terms in Expressed Topics and Initial Access Points--1A: t-Test for Two Related Samples	50
18	Number of Terms in Expressed Topics and Initial Access Points--1C: t-Test for Two Related Samples	53
19	Mean Number of Terms in Expressed Topics and Initial Access Points	55
20	Chi-Square Test of the Relationship between Type of Initial Access Point and Type of Library	59
21	Types of Scans Ranked by Frequency of Occurrence	61

LIST OF TABLES

22 Chi-Square Test of the Relationship between Number of Scans and Access Points 66

23 Types of Refinements Ranked by Frequency of Occurrence 67

24 Chi-Square Test of the Relationship between Type of Initial Access Point and Number of a priori Refinements 71

25 Chi-Square Test of the Relationship between Type of Initial Access Point and Number of Interactive Refinements 73

26 Frequency and Type of Aid Requested from the Librarian 76

27 Frequency of Bibliographic Information Recorded by 189 Subject Searchers 78

Report Number: OCLC/OPR/RR-83/1
1983 February 4

EXECUTIVE SUMMARY

A. PROJECT OBJECTIVES

The primary objective of the Subject Access Research Project was to identify features for an online public access catalog (OPAC) which support the search tactics presently employed by library users performing manual searches at the library catalog. To achieve this objective, three areas were studied:

1. Subject searcher's process of searching the library catalog (both COM and traditional card catalogs).

2. Features for an OPAC which are presently supported by subject searching of the library catalog.

3. A measure of power for a comparison of the traditional library catalog with the online public access catalog.

B. METHODOLOGY

The major tasks involved in completing the Subject Access Research Project are depicted in Figure 1. Protocol analysis was used to collect data on subject searches at the library catalog. The data collectors approached catalog users who were beginning their catalog searches, requested their participation and obtained their permission for inclusion in the study, and audiotaped users' verbalized thoughts throughout the course of their search. When searchers completed using the catalog, the data collectors obtained information about their background, education, catalog training, demographic characteristics, and the subject search itself.

Transcribed protocols were codified into a format useful for data analysis. The activity codes of Penniman (1975) were the foundation for a coding scheme devised for transcribed protocols. The coding scheme devised consisted of eight search activity codes, amplified by constituent elements which specified in detail what searchers thought or how they acted. These search activity codes are:

1. Access Point--consultation of the catalog with a word or words.

2. A priori Refinement--application of a criterion to the output of the subject search--occurs before the search of the catalog begins.

3. Scan--reading or perusal of textual information in or near the subject catalog.

4. Interactive Refinement--application of a criterion to the output of the subject search--occurs after the search of the catalog begins.

5. Aid from the librarian--consultation with the librarian sought and obtained by the searcher about the subject search.

6. Flag--searcher's mental or physical marking of a specific location in the catalog.

7. Record--mental or written note of information from the catalog.

8. Exit--physical departure of searcher from the catalog.

Search activity codes and their constituent elements are listed in detail in Table 1. A constituent element to represent a searcher's spoken thought and/or physical action was assigned whenever the searcher's protocol indicated a change in thought and/or action from one element to another. Thus, elements bear no relationship to the length of time necessary nor the extent or amount of thought or action necessary to accomplish the thought or action. They are only indicative of a change in the searcher's thought and/or action from one element to another in the course of the subject search.

C. DATA ANALYSIS

1. Participating Libraries

Data was collected at one university, three public, four high school, and four college libraries in central Ohio (Table 2). No attempt was made to collect a random sample of searchers by type of library. Libraries participating in the study reflect the types of libraries which comprise OCLC membership. Protocols of 228 searchers were audiotaped at library catalogs; 39 were eliminated from the analysis because of possible bias induced by the data collector.

2. Demographic and Background Data of Searchers

No attempt was made to obtain a random sampling of searchers; rather, searchers reflected a cross-section of the types of catalog users at OCLC member libraries.

- 72% of the searchers were ages 15 to 24. University and public library catalog searchers were more evenly distributed into higher age ranges but the 15-to-24 age range still predominated (Table 3).

- Statistically, public library catalog searchers visited the library on a weekly basis; searchers at high school, college, and university libraries visited the library on a daily basis (Table 5). The daily frequency of the latter group could be attributed to the proximity of the library and/or the need to visit the library to complete class assignments on a timely basis.

- 73% of the searchers had taken a class in which searching the library catalog was taught (Table 6).

- Subject searchers' most frequent purpose for searching the library catalog at high school, college, and university libraries was to write a paper. Their most frequent purpose for searching the library catalog at public libraries was to find personal reading materials (Table 7).

3. Findings about Frequency of Constituent Elements in Subject Searches

   - The "average" subject search has the following number of search activity codes, rounded to the nearest whole number (Table 8):

     2 access points
     2 a priori refinements
     3 scans
     0.8 interactive refinements
     0.1 aids from the librarian
     0.1 flags
     1 record
     1 exit

   - The average number of constituent elements per search was 9.73. Over two-thirds of the 189 searches in the study were assigned 10 or fewer constituent elements.

   - Subject searches containing one access point and less than the average number of constituent elements were characteristic of bookshelf browsing searches. Bookshelf browsing is an approach to subject searching in which the searcher locates a subject heading in the catalog, determines the generic class number assigned to books with that subject heading, and then goes to that class number in the bookshelves to browse through the books. A significant relationship was found between the number of access points per search and the number of constituent elements assigned to individual searches (Table 9).

   - Bookshelf browsing searches were more common approaches to search strategy for subject searchers at public, college, and high school libraries than for searchers at university libraries. A significant relationship was found between the number of constituent elements assigned to individual searches and the type of library at which the search took place (Table 10).

4. Findings about Access Points

   - Access point was defined as the consultation of the library catalog with a word or words. Approximately one of every three subject searchers consulted more than one access point (Table 12). This finding agrees with perseverance scores in previous catalog use studies (Markey, 1980).

- Searchers who persevered, regardless of the type of initial access point, varied the terminology of their expressed topic, preferring to consult an access point which was a partial (1A) or different representation (1C) of their expressed topic (Table 13).

- A greater percentage of searchers who sought assistance had more access points in their searches than the percentage of searchers who did not seek assistance. A significant relationship was found between the number of access points and the likelihood librarian assistance would be requested.

- The vocabulary of initial and subsequent access points categorized as partial representation of (1A) or different from expressed topic (1C) changed between the time the searcher approached the subject catalog and the time the searcher accessed the catalog. The number of search terms in access points which were partial (1A) or different representations (1C) of expressed topics was significantly lower than the number of search terms in expressed topics (Tables 17-18, Appendixes C3, C6).

- Subject searchers approached the library catalog with an expressed topic in mind and, at times, changed the way in which they expressed that topic when they accessed the catalog, especially by designating the access point in fewer terms than the expressed topic.

- The mean number of terms in initial and subsequent access points categorized as same representations of the expressed topic (1B or 1F) was roughly equivalent to the number of terms in initial and subsequent access points categorized as partial representations of (1A) or different from (1C) the expressed topic (Table 19). (In nearly all cases "terms" is used as meaning an individual word, rather than a group of words considered as a single item.)

- A possible reason for the change (or reduction) in terminology from expressed topic to access point during the search was that the searcher had failed in previous subject searches when consulting the exact representation of an expressed topic because its vocabulary was too detailed or specific. Searchers have thus "learned" to simplify the expressed topic by reducing it to one- or two-word general terms when accessing the catalog. The high percentage of high school students in this study whose expressed topics were exact representations of their access points (1B) contrasted with the lower percentage of college, university, and public library catalog users whose expressed topics were exact representations of their access points (Table 20). Perhaps, as high school students advance in their schooling, their expressed topics will become more specific and complex in vocabulary. However, they may continue to rely on the simple, general, brief statements that were so successful in earlier searches and transform their complex expressed topics to brief, inclusive, and/or general terminology.

5. Findings about Scans

- o A scan was defined as the reading or perusal of textual information in or near the subject catalog, e.g., guide cards, titles, or call numbers and Library of Congress Subject Headings volumes.
- o 74% of the catalog information scanned by searchers of the subject catalog was subject-rich information (e.g., titles and subject headings).
- o Not one of the 189 subject searchers consulted the Library of Congress Subject Headings or other subject heading guide in the course of a subject search.
- o Tracings of titles were scanned infrequently. Searchers' references to tracings during their audiotaped subject search (e.g., statements that "summarize the book" or "tell what the book is about") suggested that searchers did not know the purpose or value of tracings.

6. Findings about Refinements in Subject Searches

- o A refinement was defined as the application of a criterion to the output of the search which occurred before (a priori) or after (interactive) the search began.
- o A priori and interactive refinements provided the searcher with a means of including additional terms in the access point. Refinements are the searchers' method of creating a partition of the database which contains a particular phrase or term, then searching for the occurrence of another phrase or term in that partition. Refinements are analogous to the "stringsearch" of SDC's ORBIT retrieval system or the Boolean feature of Claremont College's OPAC.
- o Refinements in searches (Figure 15, 16, 17, and 25) are comparable to the Boolean operations OR, NOT, and AND for postcoordinating search vocabulary.
- o Since searchers who employed a priori and interactive refinements were often extending the vocabulary of the initial access point, such searchers usually consulted a partial (1A) or different (1C) representation of their expressed topic. Tests of a relationship between initial access point and a priori or interactive refinements (Tables 24 and 25) produced significant results which support the notion that refinements are searchers' means of extending the vocabulary of their initial access point.
- o Refinements in subject catalog searches bear similarities to "browsing" (Lancaster, 1968; Hildreth, 1982a). There were few protocols in the present study in which refinements were used to assist in browsing (Figure 17). In contrast, refinements were usually searchers' attempts to postcoordinate search terminology.

- Refinements typically considered limits (e.g., publication date, format, language) occurred infrequently in searches because of the predominance of refinements associated with search terms (Table 23). If Boolean operations are implemented in OPACs, limits might be more frequently incorporated into subject searches of online public access catalogs.

7. Findings about Aids, Flags, Records, and Exits

- Fourteen of the 189 searchers in this study sought assistance from the librarian during their search. The most frequent requests dealt with vocabulary for subject searching.

- Six of 189 searches terminated after the request for assistance from the librarian. These were instances where the librarian indicated that, for the particular case, no further search options existed within the given library location.

- Flags (i.e., the physical or mental marking of a specific location in the library catalog) occurred infrequently (15 times).

- In nearly two-thirds of the subject searches, the only catalog information recorded was the call number or call number areas.

- There were only 13 occasions (in 189 searches) when the searcher recorded full bibliographic information. This indicates that full bibliographic records of individual titles will probably be infrequently displayed by OPAC searchers, especially if brief and intermediate record displays enumerate catalog information frequently scanned (i.e., subject headings, title) and frequently recorded (i.e., call number).

D. FEATURES FOR OPACs SUGGESTED BY AD HOC PROCESS MODELS OF SUBJECT SEARCHING

1. Composite Ad Hoc Process Model

An ad hoc process model of subject searching was constructed from the analysis of 189 subject search protocols (Figure 18). It was constructed by examining the sequence of constituent elements assigned to every search. If the sequence of constituent elements of a search did not fit the model, the model was altered to accommodate the search and others that had been examined. The ad hoc process model represents the process of subject searching in a flowchart manner. Although no one search fits the composite model exactly, every search matches a portion of it.

Subsets of the composite model disclose the variety of approaches and tactics used to search the library catalog. Features for an OPAC which are presently supported by subject searching of the library catalog are suggested by subsets of the composite ad hoc process model. To support the present search tactics of library catalog searchers, such features must be implemented in online public access catalogs. Thanks to technology, the OPAC has the potential to

support features never possible in subject searching of the library catalog. Such features are also suggested by subsets of the composite model.

2. Bookshelf Browsing as an OPAC Feature

o Importance

Bookshelf browsing was a frequent approach to searching the library catalog. Characteristically, bookshelf browsing searches: 1. are assigned less than the average (10) number of constituent elements, 2. have a single access point, 3. have 0 or 1 scan, 4. are performed at high school, college, and public libraries, and 5. utilize memorized call number or call number areas. A brief four-part scenario describes bookshelf browsing at the library catalog. The searcher: 1. consults the expressed topic or some facet of it in the subject catalog, 2. scans the subject heading or call number of the first item encountered in the catalog which bears the subject heading, 3. memorizes the call number or call number area, and 4. goes to the library stack(s) which bear the call number. (See protocols, Figures 2 and 5, and ad hoc process model, Figure 19.)

o Required OPAC features

Alphabetical displays of subject headings in response to the access point input to the OPAC. In response to the access point(s) the searcher enters, the OPAC, like the library catalog, must display controlled vocabulary terms and phrases which are alphabetically adjacent or nearby the terminology the searcher enters. This alphabetical display should contain subject headings, number of postings per subject heading, and the class numbers in common to the majority of books assigned a particular subject heading (Figure 20).

o Suggested OPAC features for future implementation

(1) Online subject guide. Like the printed Subject Guide to the Collection of the Public Library of Columbus and Franklin County, the online subject guide would list the subject headings in an alphabetical display, the class number areas assigned the same subject heading, and the number of books classed in each area (Figure 21).

(2) Online classified lists or logical outlines. See the next section of this chapter, Vocabulary Aids as OPAC Features, for this suggested OPAC feature for future implementation (i.e., suggested feature (3), online classified lists or logical outlines).

3. Vocabulary Aids as OPAC Features

o Importance

Subject access points seldom exceeded two terms, but, because of the variety and complexity of the controlled vocabulary in online displays, such displays may stimulate searchers to respecify their access point utilizing the vocabulary of the online displays.

Subject searches with multiple access points were marked by searchers' inability to find the exact terminology to represent their expressed topic. (See protocols, Figures 29 and 31, and ad hoc process models, Figures 30 and 32.) Online vocabulary assistance is especially needed because OPACs may be placed in locations other than libraries (e.g., dormitories and university or city offices).

- Required OPAC features

(1) Online cross references to "approved" subject headings (e.g., MELVYL's cross-reference feature for author names). Hildreth (1982b, p. 121), underscoring the necessity for cross references in OPACs, advised that OPACs without an online cross-reference feature "deprive the users of aids that they may have already have grown accustomed to (and benefitted from) when searching the card catalog."

(2) Proximity of printed Library of Congress Subject Headings (LCSH) and reference librarian. Both the printed guide and a librarian knowledgeable and experienced in OPAC searching should be nearby or accessible to OPAC searchers.

- Suggested OPAC features for future implementation

(1) Related term displays of LCSH compiled from the syndetic structure of the printed (Figure 33) or machine-readable LCSH. Related term displays should include class numbers (generated from 053 fields of LC subject authority data and match of local library holdings) so that searchers can go directly to the bookshelves to browse items rather than display catalog information on individual items at the OPAC.

(2) Keyword-in-context displays of subject headings, which, if lengthy (e.g., more than 50 terms), categorize subject headings by subdivision (Figure 35). Postings and class number areas included in such displays would assist bookshelf browsers.

(3) Automatic display of logical outlines or classified lists to help searchers view subject vocabulary terms in context (Figure 37). The two extracts of the Dewey Decimal Classification Schedules in Figure 7 of books assigned the subject heading "Information Theory" illustrate the class number 621.38 embedded in an electrical engineering context and the class number 001.539 embedded in a communication theory context. Subject terms and phrases from schedule captions could be used to augment the subject vocabulary of bibliographic records to maximize the possibility of the searcher's matching the vocabulary of the subject catalog. Such a display of the library's classification in an online search might also aid the searcher to select the bookshelf location which best reflects his own perspective of the topic. The searcher could thus go directly to the bookshelves to browse rather than display catalog information on individual titles at the OPAC.

4. Refinements as an OPAC Feature

o Importance

Refinements were indicative of searchers' makeshift postcoordination of search terminology. A priori and interactive refinements primarily dealt with a specific term or phrase in which searchers were interested, and simulated the "stringsearch" of SDC's ORBIT retrieval system and Boolean search capability of Claremont College's OPAC. Refinements were searchers' way of creating a partition of the database containing a particular term or phrase, then searching the partition for the occurrence of another term or phrase. (See protocols, Figures 15-17, 22, and 25, and ad hoc process models, Figures 23 and 26.)

o Required OPAC features

(1) Boolean operations, allowing searchers to alter the results of a single search argument. A minimal Boolean capability should not produce irrevocable results such that the searcher has to repeat the initial search argument to "undo" the damage. Figures 27 and 28 show explicit and implicit Boolean operations of Mankato State University's MSUS/PALS and Syracuse University's SULIRS OPACs.

(2) Limit feature. Limit features may be used more often in OPACs with Boolean operations than they are in searching the traditional library catalog. Suggested criteria for a limit feature are: publication date, language, dates covered, illustrations, level of specificity, and library location.

(3) Suggestive prompts following access points with large postings. When an input access point retrieves 200 or more titles in an OPAC search, the system response should include the number of postings and a suggestive prompt which asks the searcher whether he would like to refine the retrieved items by entering additional search terminology. Likewise, when the searcher scans through displays of 100 or more items, a suggestive prompt could ask the searcher what nonrelevant terminology continually appears in the list of items retrieved, or what terminology that is particularly interesting has (or has not) appeared in the online display. Searcher's responses to suggestive prompts would activate appropriate Boolean operations which would be transparent to searchers.

o Suggested OPAC features for future implementation

Alternatives to free-text (component word) subject retrieval and Boolean operations which maintain the integrity of precoordinated subject headings are:

(1) Rotation of subject headings when producing index contents of inverted files or derived search keys. This approach is analogous to DIALOG's "Field" or "F" search which restricts retrieval to matches in a single field. If this approach is used in the creation of indexes, searchers will be able to match subject headings' innards

rather than being required to match subject headings letter-for-letter beginning with the term in the first position.

(2) Keyword-in-context displays of subject headings. This feature has already been implemented in the University of California's MELVYL (Figure 34). Keyword-in-context displays collocate subject headings with terms in common regardless of the position of input terms triggering the display in a subject heading (i.e., first position, third position, etc.).

(3) Feedback from the OPAC searcher about relevant output. CITE, a prototype OPAC of the National Library of Medicine, requests feedback from the searcher about relevant terminology in displayed items, automatically incorporates the searcher's feedback into the ongoing search strategy, retrieves additional items, and ranks output on a best-match-of-searched-terms basis.

5. Tracings as an OPAC Feature

- o Importance

Tracings were scanned infrequently by subject searchers (type 3F in Table 21). Since searchers referred to tracings as statements that "summarize the book" or "tell what the book is about," it seems doubtful that they knew either the purpose of tracings or how they could be incorporated into the ongoing search strategy.

- o Required features

(1) Include tracings in bibliographic record displays produced as a result of subject searches.

(2) Suggest the incorporation of relevant tracings in displayed bibliographic records into the ongoing search strategy. Tracings should be labeled so as to suggest that relevant tracings could be used as search terms. Such a suggestion could be introduced as a prompt following either intermediate or full bibliographic record displays or as a reminder in printed user guides.

- o Suggested OPAC features for future implementation

Include a predefined search tactic which suggests subject terminology to the searcher based on an automatic analysis of controlled vocabulary common to retrieved items. Such a feature has already been implemented in PaperChase, the OPAC of Beth Israel Hospital, to suggest Medical Subject Headings (MeSH) to the searcher based on an analysis of MeSH common to items retrieved through free-text searches of title words (Cochrane, 1982b). Through such an OPAC feature, the system, rather than the searcher, analyzes tracings. It then displays tracings which are common to the greatest number of retrieved items to the searcher who can then incorporate pertinent tracings in the search strategy.

6. Displays of Bibliographic Records as an OPAC Feature

o Importance

Subject searchers primarily scanned subject-rich catalog information (Table 21). In approximately two-thirds of the 189 searches, the call numbers or call number areas were recorded or memorized. There were only 13 searches in which full bibliographic information was recorded.

o Required OPAC features

(1) Displays of brief bibliographic records in subject searches must must include title and call number.

(2) Displays of intermediate bibliographic records must include subject headings, title, and call number (Figure 13).

7. Flags as an OPAC Feature

o Importance

Flags were defined as the physical or mental marking of a specific location in the library catalog. They occurred infrequently in subject searches (i.e., 15 times). However, the catalog itself provides the searcher with physical clues about where and what has been searched (e.g., guide cards and drawers), whereas the OPAC, especially a CRT, does not give searchers as many physical clues unless paper copy is recorded simultaneously and is available within arm's reach.

o Required OPAC features

(1) Record of subject vocabulary searched in the OPAC. Such a feature is analogous to DIALOG's "display sets" or ORBIT's "history" command.

(2) Backward and forward browse capabilities. Online vocabulary displays (e.g., alphabetic arrangements and keyword-in-context displays) and command capabilities for displaying ranges of and individual titles in brief, intermediate, and full bibliographic formats must allow searchers to move forward and backward in lists, just as library catalog searchers can presently go backward and forward by flipping catalog cards or turning a knob on a COM (computer output microform) reader.

(3) Postings information accompanying system responses to input access points. Searchers can presently estimate the number of items bearing the same subject heading by flipping blocks of catalog cards to find out where the next subject heading begins. Postings information would provide an exact count and should be incorporated into system responses to input access points.

8. Scans as an OPAC Feature

- o Importance

  Searchers rapidly scanned partial subject headings to arrive at the desired area of the subject catalog. When they arrived in the vicinity of the heading of interest, they rapidly scanned subject headings and their subdivisions. After they found the subject heading of interest, they rapidly scanned titles and their call numbers.

- o Required OPAC features

  (1) Browsing feature for subject headings. A browsing feature should permit rapid scanning of subject headings in the same vicinity of the alphabet.

  (2) Browsing feature for bibliographic records. A browsing feature should enumerate at least title and call number and allow rapid scanning of catalog information and movement from one record (regardless of format, i.e., a brief, intermediate, or full bibliographic display) to the next.

  Suggested approaches to implementing such features involve: (a) a command whose characters are mnemonic to the task, (b) a command whose characters are similar to the characters previously used to transact the same function, and (c) a suggestive prompt to inform searchers how to browse. This prompt should appear at the end of the system response which typically precedes browsing of subject headings or bibliographic records.

E. THE USER-ORIENTED INTERFACE AND POWER IN THE ONLINE PUBLIC ACCESS CATALOG

There are aspects of the online public access catalog which have never been available to searchers of the library catalog and will therefore be wholly new to those searching OPACs. Many OPAC searchers are unaware that the catalog does not physically reside in the computer terminal at their fingertips. Rather, it is electronically transmitted to and displayed on the terminal in response to searchers' messages. Presently, those messages are not only activated by searchers using a typewriter-like keyboard, function keys, or a touch-sensitive screen, but the messages must be tailored to accomplish the specific function (e.g., display a specific item, display a range of items, search for a topic, or browse subject headings) and to communicate with the particular OPAC (e.g., display 1, find su islam, brws islam). The constancy of the library catalog cannot be carried over to the online public access catalog, primarily because the physical skills inherent in library catalog searching must be translated into messages understandable to searcher and computer alike.

The constancy of the library catalog--the very quality that contributes to its ease of use, familiarity, and unimposing nature--is virtually impossible to build into the OPAC because the medium of the OPAC obviates the physical

activities inherent in library catalog searching. Thus, we cannot rely on our models of traditional catalog searching when incorporating the constancy of the library catalog into the online public access catalog. OPACs add an entirely new dimension to catalog use--the user-system interface.

Our knowledge of traditional library catalog searching cannot be relied upon when designing the user interface of OPACs because no such interface was ever present in library catalogs. Hildreth's (1982b, p. 59) "Guideline Clusters for the User-Oriented Interface" (Figure 38) can be viewed as criteria for the evaluation of the user-system interface. As new OPACs are developed and existing ones enhanced, system designers can use these guidelines in system evaluations to ensure that a user-oriented interface has been implemented, for the user-oriented interface is the crucial component of the online public access catalog, through which the searcher unlocks the system's power.

Figure 39 depicts the necessary components of OPACs with as much or more power than library catalogs. An online public access catalog in which the searcher's model of the information retrieval process of searching traditional library catalogs, be they card or COM or printed book catalogs, is implemented (and one which includes all the "required" features disclosed in this report) may have the same potential as the traditional subject catalog, but a user-oriented interface is needed to make that OPAC match the power of the traditional library catalog. In the absence of a user-oriented interface, (vs. a computer-oriented one) such an OPAC supports less than the present search tactics of subject catalog searchers because of the inherent user-oriented nature (i.e., ease of use, familiarity, and unimposing nature) of the library catalog. Likewise, an OPAC enhanced with features never before available to library catalog searchers, (e.g., title and subject keyword access and keyword-in-context vocabulary displays) may have more potential than traditional catalogs, but the inclusion of a user-oriented interface actually imbues that OPAC with more power than the traditional library catalog. The user-oriented interface is a necessary communications component of online public access catalogs, regardless of the type, number, sophistication of search features, or whether the OPAC has as much or more potential than the traditional library catalog.

Report Number: OCLC/OPR/RR-83/1
1983 February 4

I. PROJECT OBJECTIVES: A MODEL OF SUBJECT SEARCHING AT THE LIBRARY CATALOG

The objective of the Subject Access Research Project was to identify features for an automated subject retrieval system (or online public access catalog--OPAC) which support the search tactics presently employed by library users performing manual searches at the library catalog. Identification of these features is necessary to ensure that development of an online public access catalog includes at least all capabilities present in the traditional library subject catalog. The following specific areas were studied:

1. Library users' perceptions of the library catalog, their expectations, and their criteria for success.

2. The information or foreknowledge that users bring to the library catalog.

3. Subject searchers' process of searching the library catalog.

4. Features for an OPAC which are presently supported by subject searching of the library catalog.

5. A measure of power for a comparison of the traditional library catalog with the online public access catalog.

A review of catalog use studies (Markey, 1980) summarized the findings about subject catalog use and its connection to library users' perceptions, expectations, and criteria for success when searching the catalog; the information that users bring to the library catalog; and the process of subject searching in the library catalog (study areas 1-3). The review provided the groundwork and impetus to launch the present study of subject catalog use, which is unique in comparison with previous studies because it links knowledge about subject searching the library catalog to the development of the online public access catalog.

The project consisted of two methods of studying the subject catalog (i.e., focus group interviews and protocol analysis). Users' perceptions, expectations, and criteria for success when searching the library catalog were collected through focus group interviews with users (e.g., faculty and high school and college students) at public, academic, and high school libraries. The results of the analysis of searchers' remarks in focus group interviews were described elsewhere by Kaske and Sanders (1980a, 1980b); these results impact on the features of OPACs and the measure of power needed for comparison (study areas 4 and 5). The analysis of catalog searchers' protocols (i.e., spoken thoughts) as they performed a subject search in the library catalog resulted in a model of the process of searching the subject catalog (study area 3). The model discloses the features of an OPAC which are required if the OPAC is to provide at least all the capabilities now present when searching the library catalog.

This report describes the achievement of the primary objective of the project, i.e., identifying features needed for an OPAC to support the tactics presently used in catalog searching (study area 4). Such features were suggested by an ad hoc process model of the user's process of searching the library catalog and from an analysis of searchers' spoken thoughts as they performed their searches. Searchers' words and the ad hoc process model only disclose features for implementation in an OPAC which will allow searchers all the capabilities they now have using the traditional library catalog. However, thanks to technology, the OPAC has the potential to support features never possible in subject searching of the library catalog.

The OPAC also adds the user interface--a new dimension--to catalog use. A user-oriented interface is the crucial component of the online public access catalog through which the searcher unharnesses the system's power (study area 5). An OPAC in which the searcher's model of the information retrieval process of traditional catalog searching is implemented may have the same potential as the library catalog but the inclusion of a user-oriented interface is needed to make that OPAC as powerful as the traditional catalog. An OPAC enhanced with features never before available to library catalog searchers may have more potential than the traditional catalog, but the inclusion of a user-oriented interface imbues that OPAC with more power than the library catalog.

II. METHODOLOGY

A. PROTOCOL ANALYSIS FOR DATA COLLECTION

The major tasks involved in completing study areas 3-5 of the Subject Access Research Project are depicted in flowchart format in Figure 1. Protocol analysis was used to collect data on subject searches at the library catalog. This approach was used by Byers (1976) to investigate the decision-making process of library catalogers and Back (1976) to devise a model of the process of online bibliographic searching. Just as protocol analysis was employed to study these processes, the method was selected for this project to study the process of subject searching at the library catalog. Christ used this approach in three interviews (1972, pp. ix-xiv), in addition to using open-ended questions to depict student use of the library catalog and their difficulties when searching the catalog.

Data collectors participated in an all-day training session administered by Byers (see Appendix A). In the training session, basic interviewing techniques for protocol analysis were covered. Data collectors conducted practice interviews with one another. Problems encountered with data collection procedures and forms during the practice interviews were resolved by revising both the procedures and forms.

Participants in the study were recruited from four types of libraries: high school, public, college, and university. The purpose of selecting libraries by type was to ensure that those libraries included in the study represented a cross-section of the types of libraries which comprise OCLC membership (a constraint imposed on the study in the charge by OCLC to the Office of Research). Neither the number of libraries selected in the study nor the number of respondents recruited per library represents a random sample; consequently, results of the study cannot be generalized to all libraries and all library users. However, findings and case studies of subject searches are discussed according to the type of library at which they occurred so that readers can decide for themselves whether this study's conclusions about subject searching can be generalized to their own libraries.

Catalog users were recruited for participation in the study on an individual basis. Data collectors approached catalog users who were beginning their search of the library catalog and introduced themselves by saying:

> "Hello, I'm _____ from OCLC. I'm working on a survey to find out how people use the library catalog so it can be improved. I'd like to have you think out loud while you use the catalog. I'm going to use a tape recorder so I don't have to take a lot of notes; is that all right? Thank you. Go ahead."

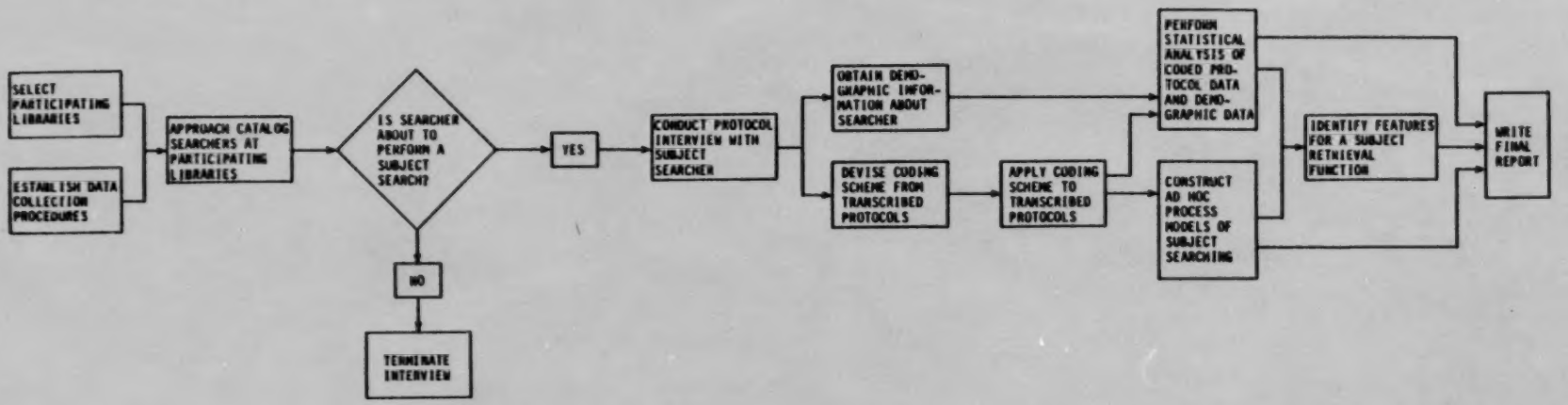

SELECT PARTICIPATING LIBRARIES
ESTABLISH DATA COLLECTION PROCEDURES
APPROACH CATALOG SEARCHERS AT PARTICIPATING LIBRARIES
IS SEARCHER ABOUT TO PERFORM A SUBJECT SEARCH?
YES
NO
TERMINATE INTERVIEW
CONDUCT PROTOCOL INTERVIEW WITH SUBJECT SEARCHER
OBTAIN DEMO-GRAPHIC INFOR-MATION ABOUT SEARCHER
DEVISE CODING SCHEME FROM TRANSCRIBED PROTOCOLS
APPLY CODING SCHEME TO TRANSCRIBED PROTOCOLS
PERFORM STATISTICAL ANALYSIS OF CODED PRO-TOCOL DATA AND DEMO-GRAPHIC DATA
CONSTRUCT AD HOC PROCESS MODELS OF SUBJECT SEARCHING
IDENTIFY FEATURES FOR A SUBJECT RETRIEVAL FUNCTION
WRITE FINAL REPORT

Figure 1. Flowchart of the Major Tasks of Study Areas 3-5 of the Subject Access Research Project

Report Number: OCLC/OPR/RR-83/1
1983 February 4

The tape recorder was then turned on and data collectors followed the catalog users through the course of their searches. If data collectors determined that the catalog users were doing something other than a subject search, the interview was terminated quickly and discretely. Data collectors followed respondents until their subject catalog search ended and taped respondents' verbalized thoughts throughout the course of their search. The data collectors asked few questions--only those intended to clarify what respondents were looking for and what text of the catalog they scanned in their searching. The interview was conducted in a nondirective manner; probes were only intended to prompt respondents to "think out loud."

When respondents completed searching the catalog, data collectors asked them questions about their background, education, catalog training, and search. Appendix B is a copy of the postsearch questionnaire which was administered orally by data collectors. Questions which data collectors could not assess themselves (e.g., the purpose of the search and the searcher's academic background) were administered orally, whereas questions which data collectors could assess themselves (e.g., the type of library and library catalog and the searcher's sex) were completed by data collectors on the basis of their own observations. At the end of each respondent's search, data collectors had obtained the following:

1. Data about the search:
    a. An audiotaped recording of respondent's spoken thoughts during a subject search of the library catalog
    b. The respondent's expressed search topic
    c. The respondent's purpose for performing the subject catalog search
    d. Recorded observations of the respondent's physical movements during the search (e.g., subject ... held these [cards] with his fingers, went back and looked them over two or three times in his selection process)
    e. Notes about the information the respondent brought to the catalog (e.g., memorized data and/or notes)
2. Demographic data about the respondent:
    a. Sex
    b. Age
    c. Years of schooling
    d. Subject field
    e. Extent of catalog instruction
    f. Frequency of library use

3. Demographic data about the library in which the respondent performed the search:

   a. Type of library

   b. Type of catalog

   c. Size of catalog

   d. Format of catalog

B. CODING SCHEME FOR ANALYSIS OF PROTOCOLS

Before the protocols could be analyzed, they had to be codified into a format from which comparisons could be made. The activity codes of Penniman (1975, p. 50) which were used in experiments to discover patterns of the online search process provided a foundation for devising a coding scheme for this study of the manual subject search process. Penniman used the following activity codes to reflect the path of information-seeking by online bibliographic searchers:

1. Begin session
2. Request database
3. Search index
4. Logic formation
5. Off-line print
6. Online display
7. Review search
8. Review commands
9. Set parameters
10. Exit database
11. End session

With these activity codes in mind, the researchers read transcripts of the audiotaped protocols and devised a coding scheme to describe the spoken thoughts and the actions of the subject catalog searchers. The coding scheme consisted of eight search activity codes, most of which were amplified by five or more constituent elements. Search activity codes and their definitions are:

1. Access Point--consultation of the library catalog with a word or words.

2. A priori Refinement--application of a criterion to the output of the subject catalog search--occurs before the search of the catalog begins.

3. Scan--reading or perusal of textual information in or near the subject catalog.

4. Interactive Refinement--application of a criterion to the output of the subject catalog search--occurs after the search of the catalog begins.

5. Aid--consultation with the librarian sought and obtained by the catalog searcher about the subject search.

6. Flag--searcher's mental or physical marking of a specific location in the library catalog.

7. Record--mental or written note of information from the catalog.

8. Exit--physical departure of the searcher from the library catalog.

Search activity codes were amplified by constituent elements which specified in detail what searchers thought or how they acted. Table 1 lists search activity codes and constituent elements. Searchers' protocols were coded into constituent elements. The assignment of a constituent element to represent a searcher's spoken thought and/or physical action was done whenever the searcher's protocol indicated a change in thought and/or action from one element to another. Thus, elements bear no relationship to the length of time nor the extent or amount of thought or action necessary to accomplish the thought or action. They are only indicative of a change in the searcher's thought and/or action from one element to another in the course of the subject search at the library catalog. In Figure 2 the coding scheme is applied to a transcribed protocol of a subject catalog search. The coded line indicates that the searcher has completed the thought or activity designated by the element and begun another thought or activity, designated by the next coded line.

Table 1. Search Activity Codes and Their Constituent Elements

SEARCH ACTIVITY CODE	CONSTITUENT ELEMENTS
1. Access Point -- consultation of the library catalog with a word or words	A. Partial representation of expressed topic found in catalog, e.g., expressed topic = "Russian Sports," access point = "Sports" B. Same as expressed topic found in catalog, e.g., expressed topic = "Information Theory," access point = "Information Theory" C. Different from expressed topic found in catalog, e.g., expressed topic = "Treatment of Blacks before the 24th Amendment," access point = "Slavery" D. Unknown, i.e., data collector was unable to determine access point E. Partial representation of expressed topic not found in catalog, i.e., same as 1A above but access point was not found F. Same as expressed topic but not found in catalog, i.e., same as 1B above but access point was not found
2. A priori Refinement -- application of a criterion to the output of the subject catalog search -- occurs before the search of the catalog begins	A. Keyword, i.e., searcher wants the keyword "Russia" to occur in catalog information. B. Key phrase C. Key concept, i.e., searcher wants the concept "Russian" to appear in catalog information, usually through the occurrence of any of a number of keywords, e.g., Russian, Soviet, Russia, U.S.S.R., etc. D. Key name and/or title E. Publication date F. Language G. Library location H. Number of pages I. Negative keyword or phrase, i.e., searcher does not want a certain word or phrase to occur in catalog information, e.g., Russian, Soviet Union, etc. J. Illustrations K. Dates covered, i.e., searcher wants the item to be about a certain chronological period, e.g., Renaissance, before 1945, etc. L. Geographic areas M. Format, i.e., searcher only wants filmstrips N. Searcher already has something similar to the retrieved item O. Number of items needed, i.e., searcher wants to retrieve a certain quantity of items P. Level of specificity, e.g., introduction to a topic, basics of a subject Q. Unknown R. Knowledge level, e.g., juvenile, for teenagers S. Time factor, i.e., searcher has a certain amount of time in mind to accomplish the task for which he needs library material T. Other

Table 1.--Continued

SEARCH ACTIVITY CODE	CONSTITUENT ELEMENTS
3. Scan -- reading or perusal of textual information in or near the subject catalog	A. Guide card B. Subject heading C. Title D. Author E. Call number F. Tracings G. "See also" reference H. Library of Congress Subject Headings I. Notes J. Full bibliographic description K. Publication date L. Library location M. Unknown N. "See" reference O. Number of pages
4. Interactive Refinement -- application of a criterion to the output of the subject catalog search -- occurs after the search of the catalog begins	A. - T. Same list of constituent elements as a priori refinements (Section 2.A. - T., this table)
5. Aid from the librarian -- consultation with the librarian sought and obtained by the catalog searcher about the subject search	(Not applicable)
6. Flag -- searcher's mental or physical marking of a specific location in the library catalog	A. Fingers, i.e., searcher sticks his fingers in the library catalog to mark a specific location B. Memory, i.e., searcher's spoken thoughts are indicative of a mental marking of a specific location in the library catalog C. Paper or other material, i.e., searcher sticks a piece of paper in the library catalog to mark a specific location D. Can't tell, i.e., the searcher flagged a catalog location but the data collector was unsure of the type of flag.

Table 1.--Continued

SEARCH ACTIVITY CODE	CONSTITUENT ELEMENTS
7. Record -- mental or written note of information from the catalog	A. Title B. Call number C. Author D. Title and call number E. Title and author F. Author, title, and call number G. Memorized call number or call number areas H. Unknown I. Full bibliographic information J. Author and call number K. Other
8. Exit -- physical departure of searcher from the library catalog	A. To stacks B. From library C. To other indexes or sources D. To study tables E. To another library F. Can't tell

Report Number: OCLC/OPR/RR-83/1
1983 February 4

DC: Excuse me, we are doing a study to see how people use the card catalog and I was wondering if you would like to participate?

S: Yeah.

DC: It is very simple. All you need to do is go through the catalog the way you normally would but think out loud for me so I can tape record what you are saying. We are trying to figure out how it is that people go through doing a search so we can make this an easier process. What is it that you were looking for before I interrupted you?

S: I am looking to see if I can find a book on analysis of financial statements.

1A

DC: OK. You found financial statements. As you are flipping through there, where is it that you are looking on the card?

S: Just. I am just looking up here to see ...

DC: The description, the second line there. OK. Does something look interesting about that?

S: About this one, I missed this one.

DC: Oh, Understanding Corporate ... Businessman's Information Guide. Guide to Understanding Financial Statements.

3C

S: Then all I do is mark down the number. A lot of times what I will do is just look for this number, it shows me where I am and then I just go back and look in the books.

7G DC: Just go back, I see. Take a look at what they have on the shelves?

S: Right. It helps.

8A DC: OK. That's it.

Key:

DC = Data Collector
S = Searcher
1A = Access Point, partial representation of topic
3C = Scan, titles
7G = Record, memorized call number or call number areas
8A = Exit, to stacks

Figure 2. Coded Protocol of a
Subject Search for Financial Statement Analysis

III. DATA ANALYSIS

A. PARTICIPATING LIBRARIES

Data was collected between November 1979 and March 1980 at one university, three public, four high school, and four college libraries in central Ohio. Although no attempt was made to collect a random sample of subject catalog searches by type of library, study participants reflect a cross-section of the types of libraries which comprise the OCLC membership, so what was learned about subject searches in public, high school, college, and university libraries can be applied to subject searches in comparable libraries throughout the United States.

During the data collection process, some data collectors felt that their prodding of searchers to think out loud might have influenced the process and/or outcome of the search. Thus, transcripts of recorded subject searches were read and checked for such a bias. There were 39 protocols in which it appeared that the data collector might have influenced the search. One of the three data collectors in the study was responsible for 33 of these 39 protocols. It was felt that her background in librarianship contributed to her influence on the search, especially her need to help when she felt that the searcher was unknowingly proceeding incorrectly in a search strategy. The 39 protocols were eliminated from the study, and the remaining 189 protocols were then coded and analyzed.

Numbers of protocols, types and names of participating libraries, and types of catalogs are listed in Table 2. Even though some protocols were dropped from the analysis, the number of protocols collected per type of library did not exceed the mean (=47) by more than 5 protocols.

B. SUBJECT SEARCHERS AT PARTICIPATING LIBRARIES

As noted, no attempt was made to obtain a random sampling of subject searchers; rather, participating libraries reflected the types of libraries in OCLC's membership. In the following discussion, the demographic and background data of subject searchers are described. Any significant characteristics of subject searchers related to the type of library in which they performed their subject searches are emphasized.

In all, 72% of the subject searchers were between the ages 15 and 24. In university and public libraries, subject searchers were more evenly distributed into higher age ranges, but the 15-to-24 age range still predominated (Table 3).

Table 2. Participating Libraries

	Type of Catalog	No. of Protocols Used in Study	No. of Protocols Deleted from Study
Public Libraries: 50 protocols			
Bexley	card-dictionary	10	4
Columbus and Franklin County	COM-dictionary	20	1
Grandview Heights	card-dictionary	20	4
High School Libraries: 46 protocols			
South	card-divided	10	4
Upper Arlington	card-divided	11	1
West	card-divided	9	0
Worthington	card-divided	16	1
College Libraries: 42 protocols			
Capital	card-divided	15	7
Denison	card-dictionary	6	0
Franklin	card-divided	20	10
Ohio Wesleyan	card-divided	1	1
University Libraries: 51 protocols (The Ohio State University Libraries)			
Engineering	card-divided	9	1
Fine Arts	card-divided	9	1
Health Sciences	card-divided	8	2
Main	card-dictionary	25	2
		189	39

Table 3. Age of Subject Searchers at Participating Libraries (N=189)

	Age: Younger than 15	15-19	20-24	25-29	30-39	40-49	Over 49	Mean age	Total
Type of Library:									
Public	1%	5%	6%	2%	8%	1%	3%	28	
High School	--	24%	--	--	--	--	--	17	
College	--	8%	11%	1%	1%	1%	--	22	
University	--	4%	14%	5%	4%	--	1%	24	
	1%	41%	31%	8%	13%	2%	4%	23	100%

Table 4. Highest Grade Completed by Subject Searchers at Participating Libraries (N=189)

	Highest grade completed: Junior High	High School	College	Graduate School	Total
Type of Library:					
Public	6%	11%	7%	2%	
High School	24%	--	--	--	
College	1%	19%	2%	1%	
University	--	21%	2%	4%	
	31%	51%	11%	7%	100%

Table 5. Frequency of Library Visits by Subject Searchers at Participating Libraries

	Frequency of Library Visits: Every n days	Every n weeks	Every n months	Every n years	Total
Type of Library:					
Public	3%	13%	8%	1.5%	
High School	10%	8%	5%	1.5%	
College	14%	5%	3%	1%	
University	17%	6%	3%	1%	
	44%	32%	19%	5%	100%

Subject searchers from public libraries varied from each other and from searchers in other types of libraries in terms of the highest level of schooling completed. Table 4 shows that, predictably, high school and college libraries are used by clusters of searchers who have attained a particular level of schooling.

Table 5 shows that subject searchers at high school, college, and university libraries are more frequent visitors of the library in which they performed their searches than are public library searchers. The latter searchers visit the library on a weekly basis in contrast to the former, most of whom visit daily. This daily frequency could be attributed to the proximity of the library and/or the need to visit the library to complete class assignments on time.

Table 6 shows that, on the average, 73% of subject searchers had taken a class in which library catalog searching was taught. Between two-thirds (university library) and four-fifths (college libraries) of searchers in all four types of libraries had been trained in searching the library catalog.

Catalog searchers expressed a variety of purposes for coming to the library. Over 70% of all searchers came to the library in connection with an educational or course requirement (e.g., to gather materials for writing a paper, to do a class assignment, or to work on research for a thesis). Table 7 lists searchers' purposes for coming to the library and the type of library in which they consulted the subject catalog. Only public libraries are marked by a different cluster of responses, namely, purposes related to personal interests (personal reading, hobbies, etc.)--an interesting but not surprising finding.

No surprises emerged in analyzing subject searchers' backgrounds. The mean age of searchers in public libraries was greater than that in other types of libraries. Public and university libraries displayed more variety in terms of the highest grade completed by searchers. High school, college, and university searchers visited the library more frequently than public library searchers. Searchers at all types of libraries were likely to have had training in library catalog searching. As subject searches are further analyzed, it is important to keep in mind that there are distinct and identifiable characteristics of public and university subject searchers.

C. FREQUENCY OF CONSTITUENT ELEMENTS IN SUBJECT SEARCHES

Transcribed protocols of 189 subject searches were coded using the scheme in Table 1. In Table 8, the frequency of search activity codes and constituent elements over all 189 searches is listed. Based on these frequencies, the "average" subject search has the following number of search activity codes, ranked in order of frequency:

- 1.57 access points
- 2.04 a priori refinements

- 2.88 scans
- 0.78 interactive refinements
- 0.09 aids from the librarian
- 0.08 flags
- 1.32 records
- 0.97 exits

for an average total of 9.73 search activity codes per search.

There are no subject searches which exactly match the number and types of search activity codes and constituent elements just enumerated, nor do these figures provide a time estimate to determine what proportion of the search was devoted to the action or thought (as designated by the element). They do indicate the extend of variation in thought and/or action during the subject search, as evidenced in Figure 3, a coded protocol of a subject search in which 11 elements have been assigned. The searcher flipped through catalog cards and scanned subject headings (3B) or titles (3C), but there is little indication how many subject headings or titles were scanned nor do her spoken thoughts indicate how long she spent scanning. However, the data collector's note indicates that this searcher's scanning of subject headings and titles was both extensive and rapid.

Figure 4 charts the frequencies of constituent elements and the number of searches assigned each frequency and graphically illustrates that the majority of subject searches were assigned few constituent elements. Over two-thirds of the 189 searches in the study were assigned ten or fewer constituent elements.

During the data collection process, it quickly became apparent that many subject searches could be characterized by a brief four-part scenario:

1. Find the expressed subject or some facet of it in the subject catalog (1=access point)
2. Scan the subject heading or call number of the first item encountered in the catalog which bears the subject heading (3=scan)
3. Memorize the call number (7G=record, call number)
4. Exit to the library stack(s) which bear the call number (8A=exit, to stacks).

Figure 5 is an example of a coded subject search which matches this scenario. The searcher in this case, like a number of other searchers in this study, checks more cards than the first card encountered to make sure that the bookshelf where he is headed contains a number of books relevant to his subject interest. The subject searcher is actually using the classified arrangement of the bookshelves to find items relevant to his search topic. Thus, the searcher's browsing of the bookshelves will actually be a

Table 6. Percentage of Subject Searchers Reporting Some Training in Searching a Library Catalog (N=177)*

	Had training in catalog use: Yes	No	Total
Type of Library:			
Public	34%	66%	
High School	20	80	
College	25	75	
University	30	70	
Average:	27%	73%	100%

*12 searchers were nonrespondents

Table 7. Purpose of Subject Searchers at Participating Libraries (N=189)

Purpose:	Type of Library: Public	High School	College	University	Total
To write a paper for example, gather materials for a paper, research thesis, or dissertation	9%	12%	12%	17%	50%
To gather class-related material for example, do homework, do class assignment, find supplemental text, do research for speech	3%	10%	6%	3%	22%
To find personal reading materials for example, for personal use, hobbies, how-to books, how-to-fix books	14%	2%	2%	5%	23%
Other for example, to kill time, regular visit, return books, look for a particular book by subject	1%	1%	2%	1%	5%
					100%

Table 8. Frequency of Search Activity Codes and Their Constituent Elements

Frequency	Percentage	Constituent elements
		1. Access Point (1.57 average per search)
103	34.7	A. Partial representation of expressed topic
74	24.9	C. Different from expressed topic
60	20.2	B. Same as expressed topic
29	9.8	F. Same as expressed topic--not found
20	6.7	E. Partial representation of expressed topic--not found
11	3.7	D. Unknown
297		
		2. A priori Refinement (2.04 average per search)
111	28.8	A. Keyword
109	28.2	B. Key phrase
36	9.3	E. Publication date
29	7.5	D. Key name and/or title
21	5.4	I. Negative keyword or phrase
13	3.4	F. Language
12	3.1	K. Dates covered
9	2.3	J. Illustrations
9	2.3	C. Key concept
8	2.1	G. Library location
8	2.1	P. Level of specificity
7	1.8	L. Geographic areas
7	1.8	R. Knowledge level
2	0.5	N. Have something similar
2	0.5	O. Number of items needed
1	0.3	H. Number of pages
1	0.3	S. Time factor
1	0.3	T. Other
0	0.0	Q. Unknown
0	0.0	M. Format
386		
		3. Scan (2.88 average per search)
316	58.0	C. Title
88	16.1	B. Subject heading
29	5.3	D. Author
22	4.0	E. Call number
19	3.5	J. Full bibliographic information
18	3.3	K. Publication date
15	2.7	G. "See also" reference
12	2.2	M. Unknown
8	1.5	F. Tracings
7	1.3	A. Guide card
6	1.1	I. Notes
3	0.6	L. Library location
2	0.4	N. "See" reference
0	0.0	O. Number of pages
0	0.0	H. Library of Congress Subject Headings
545		

Table 8.--Continued

Frequency	Percentage	Constituent Elements
		4. Interactive Refinement (0.78 average per search)
45	30.6	I. Negative keyword or phrase
42	28.5	B. Key phrase
31	21.1	A. Keyword
8	5.4	N. Have something similar
6	4.0	P. Level of specificity
5	3.4	D. Key name and/or title
2	1.4	C. Key concept
2	1.4	J. Illustrations
2	1.4	M. Format
1	0.7	E. Publication date
1	0.7	G. Library location
1	0.7	Q. Unknown
1	0.7	R. Knowledge level
0	0.0	F. Language
0	0.0	H. Number of pages
0	0.0	K. Dates covered
0	0.0	L. Geographic areas
0	0.0	O. Number of items needed
0	0.0	S. Time factor
0	0.0	T. Other
147		
17	100.0	5. Aid from the librarian (0.09 average per search)
		6. Flag (0.08 average per search)
8	53.3	B. Memory
7	46.7	D. Unknown
0	0.0	A. Fingers
0	0.0	C. Paper
15		
		7. Record (1.32 average per search)
83	33.3	B. Call number
65	26.1	H. Unknown
34	13.7	G. Memorized call number or call number areas
14	5.6	F. Author, title, and call number
13	5.2	I. Full bibliographic information
12	4.8	D. Title and call number
12	4.8	E. Title and author
11	4.5	J. Author and call number
3	1.2	A. Title
2	0.8	C. Author
0	0.0	K. Other
249		
		8. Exit (0.97 average per search)
144	78.7	A. To stacks
15	8.2	B. From library
13	7.1	C. To other indexes or sources
6	3.3	F. Can't tell
3	1.6	D. To study tables
2	1.1	E. To another library
183*		

Total number of constituent elements = 1839

*Does not equal total number of searches (N=189) because some searchers exited as a result of 5. (i.e., aid from the librarian).

DC: Hi. Excuse me. I'm doing a study to see how people use the card catalog and I was wondering if you'd be willing to participate? Basically all you need to do is go through the card catalog the way you normally would but just think out loud for me so I can tape record what you are thinking. Where are you going to start?

1A S: Christmas.

DC: All right.

S: Are we on?

DC: Yes.

2A S: Christmas. And I want particularly is Christmas in Germany. Whether it's a separate listing I don't know. I'm going to thumb very quickly to see. Yeah. Here's Christmas--Iowa, history--Fiction.

3B OK. I might check this one out. Christmas History. That's 394.

4A,7H I'm interested now, wait a minute, here's Southern States, Sweden, United States, Wales, Christmas Card, Christmas Carol. The other

3B item I might look up is Germany, Christmas. And then if I still can't get anything I'll go to one of the librarians. I'm not quite

3C,7B sure how it's going to be listed. Christmas Through the Years. 394.268. That's in the same area. The Christmas Tree. OK. I'll check that one

3C out too.

DC: Are you particularly interested in Christmas trees?

7G S: No. OK, that's it. I'll go check those out and then I'll go back to

8A look in Germany, Christmas to see if there is anything listed that way.

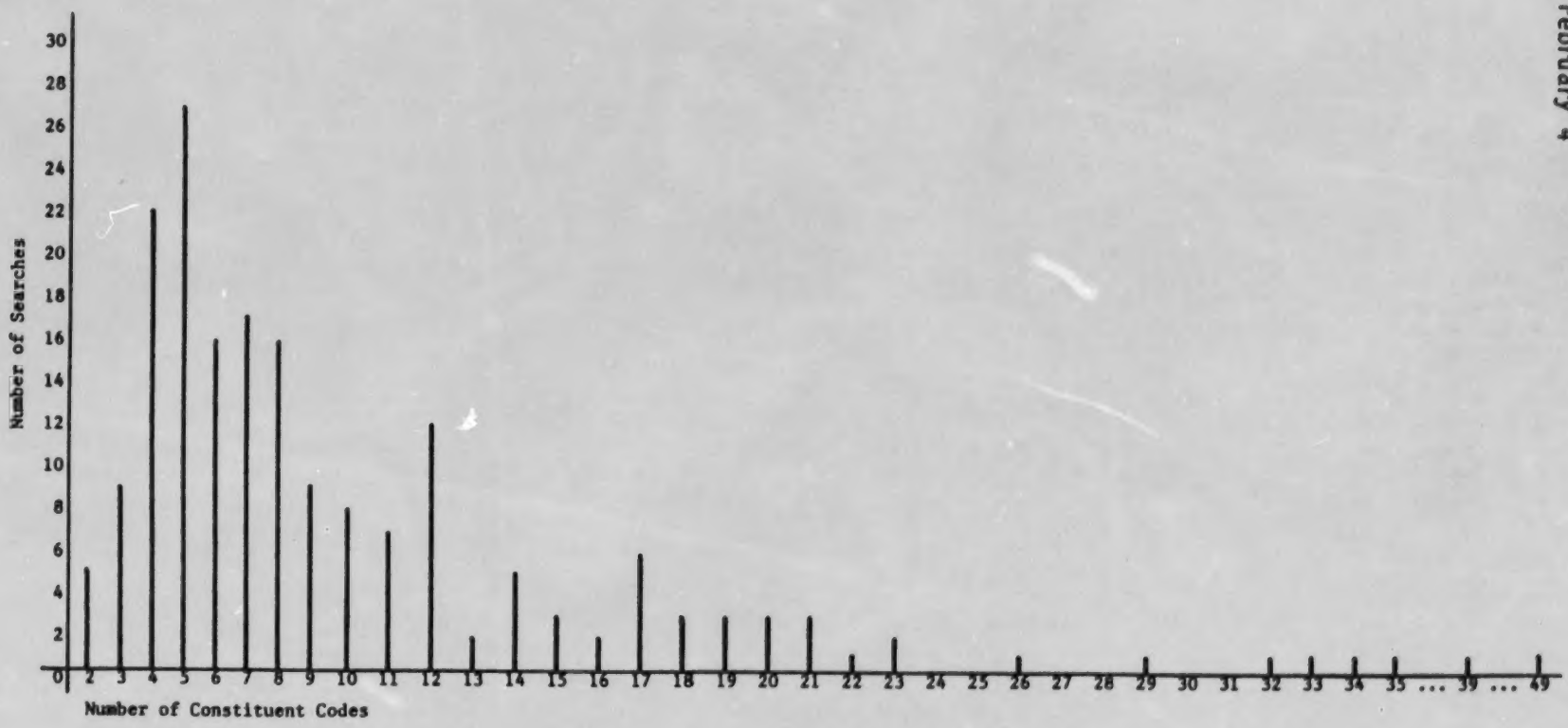
DC: Could I ask you just a couple more questions before you leave?

(DC's Notes: This person was extremely familiar with the card catalog. She whipped through it simply looking at the top at the main headings. Later on when I asked her about her background for the demographic stuff she told me that she had taught 3rd graders how to use the card catalog.)

Key:

DC = Data Collector
S = Searcher
1A = Access Point, partial representation of topic
2A = A priori Refinement, keyword
3B = Scan, subject headings
4A = Interactive Refinement, keyword
7H = Record, unknown
3B = Scan, subject headings
3C = Scan, titles
7B = Record, call number
3C = Scan, titles
7G = Record, memorized call number
8A = Exit, to stacks

Figure 3. Coded Protocol of a Subject Search for Christmas in Germany



Average number of constituent codes per search = 9.73
Median = 7
Mode = 5

Figure 4. Frequency of Constituent Codes and Number of Searches Assigned Each Frequency

S: I just checked the computer (DC's Note: a COM catalog) because I don't know the title exactly.
I am trying to find a book about information theory.
DC: Are you looking for something--a heading that says Information Theory?
1B S: Yes. Information Theory.
DC: Are you looking for a particular book you have seen before?
S: Yeah. I should check on the bookshelf because, you know, even if I
3E do see something here, there will be something different there. I
7G just go to this area and check it. These are all in the same area
8A so I can see them there. There is one more.

Key:

S = Searcher
DC = Data Collector
1B = Access Point, same as expressed topic
3E = Scan, call number
7G = Record, memorized call number
8A = Exit, to stacks

Figure 5. Coded Protocol of a Subject Search for Information Theory Assigned Four Constituent Elements

---

Information structures (Computer science) SEE Data structures (Computer science)

Information theory. 621.38(3); 001.539; 364; 501; 519.7; 651.8; SEE ALSO Automatic control; Coding theory; Data transmission systems; Information measurement; Language and languages; Machine translating; Mathematical linguistics; Modulation theory; Punched card systems; Semantics; Signal theory (Telecommunication); Statistical communication theory; Switching theory; Telecommunication

Information theory--Addresses, essays, lectures. 006

Information theory--Congresses. 519.7

Information theory in biology. 001.53

Information theory in psychology. 001.535; 301.402; SEE ALSO Human information processing

Informational measure SEE Information measurement

Informed consent (Medical law) 616.99449

NOTE: Number in parentheses following class number denotes number of books with that class number.

Figure 6. Subject Guide Printed Display of Subject Headings, Class Numbers, and Number of Books per Class Number

continuation of the subject search. Printed tools which can help subject searchers determine what area(s) of the bookshelves contain items relevant to their search topic are the library shelflist, Subject Guide, and classification schedules. Had the library where this searcher performed his subject search made its shelflist available for public access, the searcher could have searched the shelflist for the memorized call number representing his topic and scanned the book titles classed nearby to determine if there were several books of interest in the class number area.

Accessing the Subject Guide to the Collection of Columbus and Franklin County heading "Information Theory" simulates the search in Figure 5. The Guide lists subject headings and class numbers and the number of books assigned that subject heading per class number. Figure 6 shows the Guide's printed display for "Information Theory" in which the number of books assigned this subject heading are used to rank the list of class numbers. A quick scan of the Guide indicates that class number 621.38 may be the most fruitful area of the bookshelves to browse since there are three books assigned the subject heading "Information Theory" and classed into that number.

These classification schedules help put library materials into subject context. The searcher who consulted the Subject Guide (Figure 6) under "Information Theory" would probably choose to browse the bookshelf at 621.38. Had the searcher compared the captions in the classification schedules of the Dewey Decimal Classification (Dewey, 1979) for 621.38 and 001.539, he might have chosen a different bookshelf location to browse, as the former class number treats "Information Theory" from an electrical engineering perspective and the latter from a communication perspective (Figure 7). Atherton's list of "Unheeded Research Findings about Catalog Use" presented at the 1978 Library and Information Technology Association's Institute on "Closing the Catalog" refers to the value of classification schemes for enhanced subject access and context: "the classification scheme used in a library can provide subject access in the online catalog, and the words in the classification schedule captions may provide the free or natural text needed for greater access to topics in the contextual areas of the schedule, thereby providing a browsing feature in online catalogs." (Atherton, 1980, p. 109.)

Of the printed tools just discussed, only shelflist browsing is presently available as a feature or capability in online public access catalogs (OPACs). Chapter IV, Section E. 3, discusses how shelflist searching has been applied in OPACs and suggests how the Subject Guide and classification schedules could be implemented as subject access, display, and browsing features of OPACs.

Although subject searchers' browsing of the library bookshelves was not observed in this study, searchers' browsing of the bookshelves was observed and described by Tagliacozzo (1970):

> When the searcher goes to the catalog to locate a subject heading and then, instead of searching through the cards filed under it, picks up the generic call numbers of those cards and goes directly to the stacks to browse through the books, we call this a shelf-list range search. (p. IV/28.)

621.3 Electromagnetic and related branches of engineering

621.31 Generation, modification, storage, transmission of electric power
.32 Illumination and lighting
.33 Traction
.34 Magnetic engineering
.36 Applied optics (Engineering optics) and paraphotic engineering
.37 Electrical testing and measurement
➤.38 Electronic and communication engineering
.39 Other branches of electrical engineering

001.5 Cybernetics and related disciplines

001.51 Communication
.53 Cybernetics

.51 Communication
For information theory, see 001.539; communication through language, 001.54; nonlinguistic communication, 001.56
.53 Cybernetics
.532 Bionics
.533 Self-organizing systems
.534 Perception theory
.535 Artificial intelligence
.538 Decision theory
➤.539 Information theory
Including recall, uncertainty, relevance, irrelevance

NOTE: The two extracts of the Dewey Decimal Classification Schedules where books are assigned the subject heading "Information Theory" show how the class number 621.38 is embedded in an electrical engineering context and the class number 001.539 is embedded in a communication theory context. Such displays of the classification in an online search may aid the searcher to select the bookshelf location which best reflects his own perspective of the topic.

Figure 7. Extracts of the Dewey Decimal Classification Schedules Where Books Are Assigned the Subject Heading "Information Theory"

To emphasize how predominant subject searches via bookshelf browsing are, the 189 subject searches recorded at the library catalog were divided into categories according to the characteristics which exemplify bookshelf browsing searches: a single access point and nine or fewer constituent elements. Over 50% of the subject searches in the study had been assigned nine or fewer constituent elements and a single access point.

To determine if a relationship existed between subject searches categorized by number of access points and constituent elements, a chi-square test was performed. In Table 9, subject searches tallied in row 1, column 1 are akin to bookshelf browsing searches, or as Tagliacozzo named them, "shelf-list range searches." The significant result of the chi-square supports the notion that bookshelf browsing searches have identifiable characteristics and are frequently performed at the library catalog.

The number of constituent elements per subject search, one of the distinguishing characteristics of browsing searches, was matched with the type of library. Table 10 gives the frequency and percentage of occurrence of subject searches assigned nine or less or ten or more codes by type of library--public, high school, college, and university. It also discloses a significant relationship between the number of constituent elements and type of library. The magnitude of the percentage of searches assigned nine or fewer elements identifies public, high school, and college libraries as the types of libraries in which bookshelf browsing searches occur most often and in which online browsing displays of the bookshelves (e.g., through an online Subject Guide or Dewey Decimal Classification schedule) may be most appropriate.

Searches which can be categorized as bookshelf browsing searches are marked by the occurrence of a single access point and nine or fewer constituent elements. Yet not all subject searches with single access points and nine or fewer codes could be termed bookshelf browsing searches. In some searches, the searcher was not satisfied with the books classified under the subject heading and chose to consult the librarian or leave the library rather than select another access point. In addition, the small number of elements is often misleading as it does not indicate how much activity took place to accomplish the task described by the constituent element.

The coded subject search transcribed in Figure 8, especially the activity of codes 3C and 3E, demonstrates that the searcher is scanning many subject headings beginning with "R" in the COM catalog to arrive at the desired subject heading "Real Estate." When she arrives at "Real Estate," she proceeds beyond, scanning subdivisions to make sure that she will find relevant items on the bookshelves.

Rapid scanning of partial subject headings to arrive at the desired area of the catalog (e.g., r-e-a-l in Figure 8), followed by rapid scanning of the subject headings and subdivisions when the searcher has arrived at the desired area of the catalog (e.g., Real Estate--Business or Real Estate--Law in Figure 8), frequently occurs in bookshelf browsing searches. In the library catalog, the operation of flipping through catalog cards one by one is a simple physical action. A browsing feature for subject headings and/or call numbers in the OPAC which would simulate the library catalog must maintain the simplicity of the library catalog to allow searchers to rapidly scan the

Table 9. Chi-Square Test of the Relationship between Number of Access Points and Constituent Elements

Null Hypothesis: Number of access points and constituent elements in subject searches are independent.
Alternate Hypothesis: Number of access points and constituent elements in subject searches are dependent.
N = 189, 2x2 Table, .01 Level, Degree of Freedom = 1, Chi-Square $\geq$ 6.64:R

Frequency Row Percent Column Percent	Number of Constituent Elements $\leq$ 9	Number of Constituent Elements > 9	
Number of Access Points = 1	96 77% 74%	28 23% 47%	124
Number of Access Points $\geq$ 2	33 51% 26%	32 49% 53%	65
TOTAL	129	60	189

Chi-Square = 13.98*
*Rejects the null hypothesis at .01, 1 degree of freedom

Thus, there is a significant relationship between the number of access points and constituent elements in subject searches.

Table 10. Chi-Square Test of the Relationship between Number of Constituent Elements and Type of Library

Null Hypothesis: Type of library and number of constituent elements are independent.
Alternate Hypothesis: Type of library and number of constituent elements are dependent.
N = 189, 2x4 Table, .01 Level, Degrees of Freedom = 3, Chi-Square $\geq$ 11.345:R

Frequency Row Percent Column Percent	Public Libraries	High School Libraries	College Libraries	University Libraries	
Number of Constituent Elements $\leq$ 9	35 27% 70%	38 30% 83%	30 23% 71%	26 20% 51%	129
Number of Constituent Elements > 9	15 25% 30%	8 13% 17%	12 20% 29%	25 42% 49%	60
TOTAL	50	46	42	51	189

Chi-Square = 11.66*
*Rejects the null hypothesis at .01, 3 degrees of freedom

Thus, there is a significant relationship between the number of constituent elements in subject searches and type of library.

S: Well, I'm looking for a real estate book so I'm looking at the R's. First, fast to "Real Estate." Wrong way.
DC: OK. What are you looking at?
S: I'm looking at "real" ... right now I'm looking for "Real Estate."
DC: Are you looking at the columns?
1B, 3C, 3E, 7B, 8A S: At the top, up there. That's r-e-a-d still. Here we are: "Real Estate," "Business," "Principles and Practices," "Law." OK. The numbers are 333, 346. I just want to get a general idea where they are. So I'll just look at 333 and 346.

Key:

S = Searcher
DC = Data Collector
1B = Access Point, same as expressed search topic
3C = Scan, subject headings
3E = Scan, call number
7B = Record, call number
8A = Exit, to stacks

Figure 8. Coded Protocol of a Bookshelf Browsing Search for Real Estate

catalog information. This could be done by using a command whose characters are mnemonic to the task and/or similar to characters previously input if the operation is a repetitive one. Another approach would be to incorporate a suggestive prompt at the end of all browsing displays which would inform the searcher how to continue the display (e.g., "FOR MORE, PRESS + OR -").

The simplicity with which scanning of catalog information is done in the library catalog is a requirement for OPACs. This discussion on the frequency of constituent codes notes the identifiable characteristics of a popular subject search called bookshelf browsing or shelf-list range search. It also points out that the small number of elements occurring in bookshelf browsing searches misrepresents the amount of scanning activity which transpires. Enhancements to subject searches are suggested for implementation in OPACs so that the searcher can actually browse the bookshelves at the OPAC. Because of the frequency of browsing searches in public, high school, and college library catalogs, an online Subject Guide or Dewey Decimal Classification schedule are suggested as appropriate enhancements to such OPACs.

D. ACCESS POINTS IN SUBJECT SEARCHES

In the present study, both the vocabulary of searchers' expressed topics and the initial access points were recorded. The expressed topic was the vocabulary used by the searcher to state what he was looking for. Access point was defined as the consultation of the library catalog with a word or words. Access points were assigned to six categories depending upon their relationship to the expressed topic:

1A. Partial representation of expressed topic--found in catalog.
Example: The searcher phrased his expressed topic as "Russian Sports" but consulted the library catalog at "Sports".

1B. Same as expressed topic--found in catalog.
Example: The searcher phrased his expressed topic as "Information Theory" and consulted the library catalog at "Information theory".

1C. Different from expressed topic--found in catalog.
Example: The searcher phrased his expressed topic as "Treatment of Blacks before the 24th Amendment" but consulted the library catalog at "Slavery".

1D. Unknown.
Example: The searcher phrased his expressed topic as "Crockpot Cookery" but the data collector could not see, did not record, etc., what the searcher consulted at the library catalog.

1E. Partial representation of expressed topic--not found in catalog.
Example: The searcher phrased his expressed topic as "Effects of Snuff" but consulted the library catalog at "Snuff" and did not find the word(s) of the access point in the catalog.

1F. Same as expressed topic--not found in catalog.
Example: The searcher phrased his expressed topic as "Female Impersonators" and consulted the library catalog at "Female impersonators" but did not find the exact word(s) of the access point in the catalog.

In Table 11 the frequency of the six types of access points is summarized and the percentage of access points found and not found in the catalog is tallied. The occurrence of nearly 300 access points in 189 searches implies that one of every two subject searchers consulted more than one access point. This summary is, however, a misrepresentation of subject searchers' perseverance since some searchers consulted as many as seven access points in a single search. Table 12 lists the number of searches and access points in subject searches. Nonetheless, two of every three searchers (124 of 189 searchers) did not consult more than one access point. This agrees with findings of four previous catalog use studies which reported that over 50% of catalog users consulted a single subject heading (Markey, 1980, p. 22).

The perseverance of subject searchers varies considerably depending on the level of success met with using the initial access point. As Table 13 shows, searchers who were unable to find their access point in the catalog (codes 1E or 1F) were most likely to persevere and consult another access point. Searchers least likely to consult a second access point were those who initially consulted the exact (1B) or partial (1A) representation of their expressed topic and found it in the catalog.

The type of initial access point is thus an indication of whether the searcher will continue the subject search by consulting the catalog at another access point. Table 14 shows the results of a chi-square test to determine whether there is a relationship between the type and number of initial access points in subject searches. The frequencies show that, in searches containing two or more access points, the initial access points probably were not located in the subject catalog (i.e., searches categorized as 1E or 1F). In contrast, few searchers whose initial access point was the exact representation of their expressed topic bothered to access the catalog more than once. Only five searchers who initially located the exact representation of their expressed topic continued to access the catalog (see Table 14). Those who continued, regardless of the type of initial access point, varied the terminology of their expressed topic, preferring 1A or 1E (partial representations) or 1C (representation different from topic).

Percentages of access points at the initial, middle, and last access point are enumerated by type in Table 15. These figures imply that searchers' consultation(s) of the subject catalog subsequent to their initial access point deviate from the terminology of the expressed topic.

Since subject searchers who consult the catalog at more than one access point have usually failed to find their initial access point in the catalog, if they alter their access point and repeatedly fail to match their terminology with that of the catalog, they may become frustrated and seek help. Table 16 shows a significant relationship between the number of access points in a subject search and the seeking of assistance from the librarian. Although only 14 of 189 searchers in the present study obtained assistance from the librarian, the number of those consulting more than one access point and obtaining assistance

Table 11. Frequency of Access Points in Subject Searches

1 Access Points:		Frequency	Percentage
1A	Partial representation of expressed topic-found	103	34.7
1B	Same as expressed topic-found	60	20.2
1C	Different from expressed topic-found	74	25.0
1D	Unknown	11	3.7
1E	Partial representation of expressed topic-not found	20	6.7
1F	Same as expressed topic-not found	29	9.7
	TOTAL =	297	100.0%

Percent of Access Points Not Found in Catalog = 16.4%
Percent of Access Points Found in Catalog = 79.9%
Unknown = 3.7%

Table 12. Number of Searchers and Number of Access Points

Number of Access Points	Number of Searchers	
1	124	65.6% of subject searchers did not consult more than one access point.
2	40	34.4% of subject searchers consulted more than one access point.
3	14	
4	7	
5	2	
6	1	
7	1	

Table 13. Perseverance and Type of Initial Access Point

Initial Access Point:	No. of searches with > 1 access point in which this was initial access point	Total no. of searches in which this was initial access point	Percentage of searchers who persevered
1A Partial representation-found	17	72	24%
1B Same as expressed topic-found	5	54	9%
1C Different from expressed topic-found	12	22	55%
1D Unknown	3	6	50%
1E Partial representation-not found	8	9	89%
1F Same as expressed topic-not found	20	26	77%

Table 14. Chi-Square Test of the Relationship between Type and Number of Initial Access Points

Null Hypothesis: Type and number of initial access points are independent.
Alternate Hypothesis: Type and number of initial access points are dependent.
N = 189, 2x6 Table, .01 Level, Degrees of Freedom = 5, Chi-Square $\geq$ 15.086:R

Frequency Row Percent Column Percent	Number of Access Points = 1	Number of Access Points = 2 or more	
Type of Access Point:			
1A Partial representation-found	55 76% 44%	17 24% 26%	72
1B Same as expressed topic-found	49 91% 40%	5 9% 8%	54
1C Different from expressed topic-found	10 45% 8%	12 55% 18%	22
1D Unknown	3 50% 2%	3 50% 5%	6
1E Partial representation not-found	1 11% 1%	8 89% 12%	9
1F Same as expressed topic not-found	6 23% 5%	20 77% 31%	26
TOTAL	124	65	189

Chi-Square = 56.09*
*Rejects the null hypothesis at .01, 5 degrees of freedom

Thus, there is a significant relationship between the number and type of access points in subject searches.

Table 15. Percentage and Type of Access Points Occurring as the Initial, Middle, and Last Access Points

Type of Access Points:	Percentage of Initial Access Points	Percentage of Middle Access Points	Percentage of Last Access Points
1A Partial representation-found	38.1	18.6	35.4
1B Same as expressed topic-found	28.6	2.3	7.7
1C Different from topic-found	11.6	46.5	49.2
1D Unknown	3.2	7.0	3.1
1E Partial representation not-found	4.7	18.6	4.6
1F Same as expressed topic not-found	13.8	7.0	0.0
Total number of searches in category	N=189	N=43	N=65

Table 16. Chi-Square Test of the Relationship between Number of Access Points and Request for Librarian Assistance

Null Hypothesis: The number of access points in subject searches and searcher's request for librarian assistance are independent.
Alternate Hypothesis: The number of access points in subject searches and searcher's request for librarian assistance are dependent.
N = 189, 2x2 Table, .01 Level, Degree of Freedom = 1, Chi-Square $\geq$ 6.64:R

Frequency Row Percent Column Percent	No. of searches in which librarian's assistance not sought	No. of searches in which librarian's assistance obtained	
Number of Access Points = 1	120 96.8% 68.6%	4 3.2% 28.6%	124
Number of Access Points $\geq$ 2	55 84.6% 31.4%	10 15.4% 71.4%	65
TOTAL	175	14	189

Chi-Square = 7.52*
*Rejects the null hypothesis at .01, 1 degree of freedom

Thus, there is a significant relationship between the number of access points and request for librarian assistance.

was significantly higher than the number of searchers who consulted only one access point and then obtained assistance.

In a previous catalog use study (Coblenz and Tagliacozzo, 1970, IV/119), 75.5% of subject catalog searchers were successful in matching the vocabulary of their initial access point with that of the catalog. In the present study, searchers' initial access points categorized as 1A, 1B, and 1C were found in the subject catalog (Table 13, column 3). The number of searches whose initial access points were categorized in this way was 72, 54, and 22, respectively. Thus, in 148 of 189 searches, the vocabulary of searchers' initial access points matched the catalog's controlled vocabulary. It appears, then, that 78% of subject searchers, a figure comparable to that reported by Coblenz and Tagliacozzo in a similar study, were successful in matching the vocabulary of their initial access point with that of the catalog. The surprising magnitude of success reported in both studies suggests that closer scrutiny of searchers' terminology in this study might prove instructive.

The vocabulary of access points categorized as 1A, 1C, or 1E changed between the time the searcher approached the subject catalog and the time the searcher accessed the subject catalog. The description of these three types of access points gives an indication of how the vocabulary of the expressed topic changed. Examples of the vocabulary of expressed topics and corresponding type-1A access points are:

Expressed Topic	Initial Access Point (1A)
Commerce and Geography of Poland	Poland
Psychological Aspects of Rape	Rape
Civil War Medicines	Civil War
World War II Tanks	World War, 1939-1945
Trance Dances of the Holy Ghost People	Trances
Description of Florida's East Coast	Florida--Description and travel

These examples show that subject searchers' initial access points are comprised of some, but not all, of the terms which make up the expressed topic. Cases where terms not in the expressed topic are represented in the initial access point (e.g., Florida--Description and travel) may be attributed to the searcher's knowledge and acceptance of a precoordinated controlled vocabulary used to represent his expressed topic. A glance at these examples discloses that the number of terms in the expressed topic exceeds the number of terms in the initial access point. The terms of expressed topics and corresponding initial access points were listed and counted to determine if significantly more terms were represented in expressed topics than in initial access points when the latter were a partial representation of the expressed topic. Appendix C1 lists the terms and number of terms in expressed topics and initial access points (1A). The results of a t-test for two related samples are given in Table 17 and confirm that significantly greater numbers

of terms are represented in expressed topics than in initial access points (1A).

This significant reduction in the number of terms in initial access points is not characteristic of initial access points only. Access points (1A) occurring after the initial access point (i.e., those that were in the middle or the last access point) were listed and their terms were counted (see Appendix C2). A t-test for two related samples was calculated; the number of terms in expressed topics was significantly greater than the number of terms in access points (1A) in the middle or at the end of subject searches (see Appendix C3). However, the number of terms in access points (1A) occurring after the initial access point sometimes exceeded the number of terms in the expressed topic, a phenomenon which did not occur in initial access points (1A).

Initial access points as partial representation (1A) of the expressed topic occurred in 38% of subject catalog searches and ranked as the most preferred approach overall (see Table 11) and initially (see Table 13). Since the searcher "watered down" his expressed topic into fewer terms, he thereby "broadened" the search and might have had to "wade through" many items which treated the topic in a general way. Only one of four searchers who accessed the catalog with an access point which was a partial representation (1A) of the expressed topic persevered and chose more than one access point (see Table 12). Figure 9 is an example of a subject search in which the searcher described his expressed topic as "Canadian Anthem," accessed the catalog at "Canada," and, rather than browse the catalog for books which directly refer to "Anthems" or find the correct subject heading "National Songs," was content to memorize the call number representing Canada and look through books to find something which refers to Canada's Anthem.

There were also access points whose terminology was different from that of the expressed topic (1C):

Expressed Topic	Initial Access Point (1C)
Treatment of Blacks before 24th Amendment	Slavery
Breakfast Goodness	Nutrition
Player Piano	Keyboard instruments
Togas	Costumes, Ancient or Roman
Boat People	Vietnamese
Frank Stella and Kenneth Noland	Abstract expressionism

Searchers who consult an access point (1C) which is different from the expressed topic run the risk of finding much library material covering the topic of interest along with many other topics. For example, a book about keyboard instruments will probably include information on player pianos as well as harpsicords, organs, clavichords, and other keyboard instruments. It

Table 17. Number of Terms in Expressed Topics and Initial Access Points--1A: t-Test for Two Related Samples

Null Hypothesis: Mean of the number of terms in expressed topics is equal to the mean of the number of terms in initial access points--1A.
Alternate Hypothesis: Mean of the number of terms in expressed topics is greater than the mean of the number of terms in initial access points--1A.

t-test for related samples, N=72
.005 level, one-tailed, degrees of freedom = 71
R: $t \geq 2.66$

$\overline{D} = 126/72 = 1.75$

$$\Sigma d^2 = 316 - \frac{126^2}{72} = 95.5$$

$$S_{\overline{D}} = \sqrt{\frac{95.5}{72(72-1)}} = .1367$$

$$t = 1.75/.1367 = 12.8^*$$

*Rejects the null hypothesis at .005, 71 degrees of freedom

Thus, the number of terms in expressed topics is significantly greater than the number of terms in initial access points--1A.

DC: Can I ask you what you are looking for?
S: Yes. I'm looking for songs on Canada -- Canadian Anthem. I'm trying to find it under Canada.
1A DC: OK. You're looking for something under Canada now?
S: Right.
DC: What are you reading?
S: J9N7. Find it in that section.
DC: So you've learned the number and now you're going to look for it?
S: Yes, I know what section it is in so I'll just go look and I might not find that but I might find what I'm looking for.
DC: So you just looked to see if you can find where it's located on the shelf?
S: Right.

Key:

DC = Data Collector
S = Searcher
1A = Access Point, partial representation of topic

Figure 9. Partially Coded Protocol of a Subject Search for Canada's Anthem (Initial Access Point Is a Partial Representation of Expressed Topic--1A)

is not as clear in these examples as it was in examples of initial access points which were partial representations of the expressed topic (1A) that the number of terms in initial access points which are different from the expressed topic (1C) is comparatively less than the number of terms in expressed topics. Appendix C4 lists the terms and numbers of terms in expressed topics and initial access points (1C). The results of a t-test for two related samples, given in Table 18, confirm that significantly more terms are represented in expressed topics than in initial access points which are different from the expressed topic (1C). The expressed topics and corresponding middle or last access points (1C) and the number of terms in each are listed in Appendix C5. The significance in the greater number of terms used to express the former as compared with the latter was established in the calculation of a t-test for two related samples (see Appendix C6). Thus, both initial access points that were different (1C) from the expressed topic and middle or last access points that were different (1C) from the expressed topic contained significantly fewer terms than the corresponding expressed topics.

The protocol in Figure 10 discloses that the searcher approached with the expressed topic "How Good a Breakfast Is" and initially accessed the catalog under "Nutrition." Understandably, books assigned the subject heading "Nutrition" would cover topics besides breakfasts. The searcher implies in the protocol that the topic will probably be contained in part in the books found in the Nutrition classification area (i.e., 613).

Initial access points which were the exact representation of the expressed topics (1B) demonstrate their similarity in generality and number of terms to the "watered down" access points of 1A and 1C (see Appendix C7). Examples of these initial access points are:

Expressed Topic and Initial Access Point (1B)

Prostitution

Anorexia Nervosa

Hypertension

Calligraphy

Biofeedback

Prohibition

Mythology

These examples, which are both expressed topics and access points (1B), are made up of one or two terms like the access points which were categorized into 1A (partial representation of expressed topic) or 1C (different from expressed topic). In contrast, the expressed topics of 1A and 1C access points were composed of two or more terms. Table 19 lists the mean number of terms in expressed topics and corresponding initial access points. Rows are boxed to emphasize how few terms made up the expressed topics when the initial access point was the same as the expressed topic (1B).

Table 18. Number of Terms in Expressed Topics and Initial Access Points--1C: t-Test for Two Related Samples

Null Hypothesis: Mean number of terms in expressed topics is equal to the mean of the number of terms in initial access points--1C.
Alternate Hypothesis: Mean number of terms in expressed topics is greater than the mean of the number of terms in initial access points--1C.

t-test for related samples, N=22
.005 level, one-tailed, degrees of freedom = 21
R: $t \geq 2.831$

$\overline{D} = 1.0455$

$$\Sigma d^2 = 69 - \frac{23^2}{22} = 44.9545$$

$$S_{\overline{D}} = \sqrt{\frac{44.9545}{22(22-1)}} = .3119$$

$$t = 1.0455/.3119 = 3.3516*$$

*Rejects the null hypothesis at .005, 21 degrees of freedom

Thus, the number of terms in expressed topics is significantly greater than the number of terms in initial access points--1C.

1C DC: You are looking for something under Nutrition.
S: I was just going to get some books to see if they had anything on how good a breakfast is.
DC: Oh, OK. So you want to evaluate a breakfast.
S: I am trying to write a persuasive paper on how good it is. I just wanted to see if they had anything so I am going to look under these really.
DC: Are you going to take down one of the numbers?
S: No. I will probably take a couple of them, maybe I'll look around 613.
DC: How do you tell if it is something you want?
S: I just read ....
DC: ...the title part of it?
S: Yeah, see if it might pertain to what I was thinking of.
DC: Do you want any that say 'Tables"?
S: No, I don't. See--About Your Diet--that may be good. I don't really need too many.
...There is an awful lot of them.
DC: Yeah. There are.
S: I think I will go in the back and find one from there. I think that's it.

Key:

DC = Data Collector
S = Searcher
1C = Access Point, different from expressed topic (i.e., "Nutrition")

Figure 10. Partially Coded Protocol of a Subject Search for How Good a Breakfast Is (Initial Access Point Is Different from Expressed Topic--1C)

Table 19. Mean Number of Terms in Expressed Topics and Initial Access Points

Type of Initial Access Point:	Mean Terms in Expressed Topic	Mean Terms in Initial Access Point
1A Partial representation of expressed topic--found	3.10	1.35
1B Same as expressed topic--found	1.46	1.46
1C Different from expressed topic	2.64	1.59
1E Partial representation of expressed topic--not found	2.44	1.22
1F Same as expressed topic--not found	1.72	1.72

Figure 11 is an example of a protocol in which the searcher's initial access point matched his expressed topic (1B). The searcher is looking for books on "Motorcycles" and, having found his initial access point in the catalog, spends the remainder of the search determining which of many relevant books he will select from the bookshelves. In contrast to subject search examples in which the initial access points were partial representations of (1A) or different from (1C) the expressed topic, the search in Figure 11 demonstrates that the searcher is using the subject information which is found on catalog cards to screen books for relevance and compare them to one another. The book selection process of this searcher is probably complete by the time he leaves the card catalog and his perusal of the bookshelves will be limited to finding the call numbers of the individual books he earlier deemed relevant.

In this subject search for books on motorcycles, the searcher knows exactly what books he will select from the bookshelves before he even arrives there. Previous subject searches for "Canadian Anthem" (Figure 9) and "How Good a Breakfast Is" (Figure 10) showed that the searcher would carry out the book selection process at the bookshelves by scanning books' innards (contents, indexes, introduction, bookjacket, etc.). In one case, the library catalog is a tool for the location of subjects and book selection; in the other searches, the catalog is only a tool for the location of subjects and book selection is a separate process that occurs after the searcher arrives at the bookshelves.

Searchers whose expressed topic matches their initial access point also use the library catalog exclusively as a tool for the location of subjects. Figure 12, a search for "Mondrian," in which the searcher's expressed topic matched his access point, indicates that the searcher will browse the bookshelves by the searcher's remark "I just take down this thing here, whatever it is, the call number .... [but] if they are all in the same area, I don't bother." This searcher appears to have greater confidence that the books located in the same class number area refer to his expressed topic than does the searcher for "How Good a Breakfast Is," because the catalog's vocabulary in the search of the former exactly matches his expressed topic and initial access point.

The evidence presented shows that subject searchers approach the card catalog with a topic in mind and, at times, change that topic when they access the catalog. In addition, the percentages in the columns of Table 20 show that high school students (the youngest searchers in this study) generally access the catalog with the same vocabulary as their expressed topic.

When the vocabulary or the access point is different from the expressed topic (1C) or a partial representation (1A or 1E) of the expressed topic, the access point is designated by fewer terms than the expressed topic. When searchers approach and access the catalog with the same vocabulary (1B or 1F), the number of terms in the expressed topic is roughly equal to the number of terms in the access points of searches that have undergone the change (or reduction) process.

One possible reason for this change (or reduction) in terminology in the course from expressed topic to access point is that the searcher has experienced failure in previous subject searches when consulting the exact representation of his expressed topic because its vocabulary was too detailed

DC: What are you looking for?
1B — S: Motorcycles. Just in general. Some information on dirt bikes and just a little bit on street bikes too. I used to race last year so I know pretty much.
DC: What part of the card are you reading?
S: The bottom part.
DC: The part where it says the title?
S: Yeah. "Motorcycle Racing." The title. I don't care about this. This would be a good one because it sounds like a general motorcycle handbook.
DC: OK. Still reading the titles?
S: Yeah. I'm just like going through this little bit seeing if that book's in reserve I can't take that out. I could have used this one too. It talks about maintenance. This one's Motorcycle Riding. I'm looking for something that would be really thrilling to talk about in a speech.
DC: Something with a lot of excitement?
S: Yeah. Something in here like Evil Knievel or something like that about all his stunts and everything. I'll just take one out that's really good and talk about it. Ought to be pretty good. Let's see. I need one more. I'll make my topic broader I guess.
DC: Make it broader by picking something ....
S: Yeah. Something else that I'd like to talk about. They only have two books and that's not enough to do a speech on. See, here's what I was looking for, "Motorcycle Racing." They talk about Arley Dirt Racing. "Motorcycle Handbook." OK, I got all three of them.
DC: So now what are you doing? Memorizing the numbers?
S: No, I'm just kind of thinking what order like if I could cancel one of these out for motorcycle racing for another one then I might just talk about dirt racing or something like that and it would be a lot easier for me to do. That's the only one they have so I'll have to take that. And "Motorcycle Maintenance." There's one more here. I lost it. "Motorcycle Handling," "Motorcycle Racing," where did that thing go? Oh, here we go.
DC: Have you read any of those before?
S: No. I learned how to ride a bike by trial and error.
DC: Just copying down the call numbers?
S: Yes.
DC: The stuff in the upper left hand corner?
S: Yes. Well, that's it I guess.
DC: That's it. (DC's Note: Subject went through all the cards and narrowed it down to three or four that he wanted and held these with his fingers, went back and looked them over two or three times in his selection process.)

Key:

DC = Data Collector
S = Searcher
1B = Access Point, same as expressed search topic

Figure 11. Partially Coded Protocol of Subject Search for Motorcycles (Initial Access Point Is the Same as Expressed Topic--1B)

DC: What is it that you are going to look under?
S: Mondrian. I start looking at the subject first, usually.
1B— DC: As you are flipping through there, what area on the card are you looking at?
S: Up here.
DC: At the top?
S: Yeah. Looking to find something in English.
DC: It looks like most of them are in French.
S: Yeah. I'll just have to settle for that then. I just take this thing here, whatever it is, the "call number."
DC: So you are just basically going to write the call numbers down of all the ....
S: What I think will be some and, especially if they are all in the same area, I don't bother.

Key:

DC = Data Collector
S = Searcher
1B = Access Point, same as expressed search topic

Figure 12. Partially Coded Protocol of a Subject Search for Mondrian (Access Point Is the Same as Expressed Topic--1B)

Table 20. Chi-Square Test of the Relationship between Type of Initial Access Point and Type of Library

Null Hypothesis: Type of access point and type of library are independent.
Alternate Hypothesis: Type of access point and type of library are dependent.
N = 148, 3x4 Table, .01 Level, Degrees of Freedom = 6, Chi-Square $\geq$ 16.812:R.

Frequency Row Percent Column Percent	Type of Library:				
	High School	Public	College	University	
Type of Initial Access Point:					
1A or 1E, Partial representation of expressed topic	7 10% 20%	18 25% 43%	18 25% 62%	29 40% 69%	72
1B or 1F, Same as expressed topic	21 39% 60%	20 37% 48%	5 9% 17%	8 15% 19%	54
1C, Different from expressed topic	7 32% 20%	4 18% 9%	6 27% 21%	5 23% 12%	22
TOTAL	35	42	29	42	148

Chi-Square = 26.37*
*Rejects the null hypothesis at .01, 6 degrees of freedom

Thus, type of initial access point and type of library are dependent.

or specific. Searchers have thus "learned" to simplify their expressed topic by reducing it to one or two general terms when they access the catalog. The results of a statistical test to determine a relationship between type of library and type of initial access point support this notion (Table 20).

Truly, a longitudinal study of subject catalog searchers would be necessary to confirm this conclusion, but the tallies in Table 20 for initial access points that are the same as expressed topics seem to indicate that high school students, the youngest and perhaps the least experienced of the library catalog searchers, prefer this approach to searching the catalog. Subject searchers in college, university, and public libraries approach the catalog with more verbose expressed topics than high school students, but may simplify, generalize, or abbreviate their initial access points in ways that render access points similar to the expressed topics of high school students. On one hand, high school students' search topics may be simple, general, and brief; on the other hand, as these high schoolers advance in the course of their educational careers, their expressed topics may become more specific and complex in terms and facets, but they may continue to rely on the simple, general, brief statements that were so successful in their early subject catalog searches and translate their expressed topics into brief, inclusive and/or general terms.

E. SCANS IN SUBJECT SEARCHES

The term "scan" refers to the reading or perusal of textual information in or near the subject catalog. In only 26 searches were no scans recorded. Table 21 summarizes the number and type of scans which occurred in 189 searches. Fewer than 25% of the subject searches involved more than three scans. In Table 21, the frequencies of scan types indicate that subject searchers overwhelmingly scan subject-rich information found on catalog cards (i.e., subject headings and titles). These two types of information accounted for 74% of the information scanned by subject searchers. However, some textual information scanned by searchers is not always present for every cataloged item (e.g., notes, guide cards, and publication date); thus, it is not surprising that the highest ranked information (i.e., titles, subject headings, and author) is usually enumerated for every bibliographic record in the library catalog.

Not one of 189 subject searchers consulted the Library of Congress Subject Headings or any other subject heading guide during the course of their subject search at the catalog. Searchers scanned "see" and "see also" references very few times, but this may have been due to the catalog's failure to contain these two helpful pointers, not to the searcher's failure to read them.

Tracings were scanned very infrequently. Because searchers who referred to tracings during their search described them as statements that "summarize the book" or "tell what the book is about," it is doubtful whether searchers knew that such statements were the controlled vocabulary which was used by the library to describe books' intellectual contents and could thus use these statements and incorporate them into their search strategy as access points to the subject catalog.

Table 21. Types of Scans Ranked by Frequency of Occurrence

Type of Scan	Frequency	Percentage of Total
3C Titles	316	58.0
3B Subject Headings	88	16.1
3D Author	29	5.3
3E Call Number	22	4.0
3J Full bibliographic information	19	3.5
3K Publication date	18	3.3
3G "See also" reference	15	2.7
3M Unknown	12	2.2
3F Tracings	8	1.5
3A Guide cards	7	1.3
3I Notes	6	1.1
3L Library location	3	0.6
3N "See" reference	2	0.4
3O Number of pages	0	0.0
3H Library of Congress Subject Headings	0	0.0
TOTAL	545	100.0%

Since subject catalog searchers are prone to scan subject-rich information, it is recommended that subject headings and titles be included in intermediate subject displays of individual items in OPACs. Tracings should be included in intermediate displays and labeled so as to suggest to the searcher that he should incorporate this vocabulary into his search strategy to obtain similar library materials. Call numbers should also be included at this display level to accommodate searchers who wish to terminate their OPAC search and browse the bookshelf. Figure 13 is an example of such a suggested display. It also shows how an OPAC can provide a helpful prompt immediately following the system's response to a search argument, telling the user how to obtain abbreviated or intermediate displays of individual items. Abbreviated displays should produce the call number and title; intermediate displays should include call number, title, and subject headings. Call number is a requirement for any and all item displays produced as a result of subject searches because of the predominance of searchers who rely on bookshelf browsing for book selection.

Examples of scans in a subject search are designated in Figure 14. The dialogue indicates that the data collector had to prompt the searcher to tell where and what she was looking at in the catalog. Interestingly, the scans in this case seem to be precipitated by the initiation of an access point, i.e., the more access points, the more scans. The converse likewise appears true. In a test of the relationship between the number of scans and access points in subject searches (reported in Table 22), a significant relationship was found. A large percentage of searches containing a single access point and zero or one scan are probably indicative of what happens in bookshelf browsing searches. We have already shown that the number of access points and constituent elements are minimal in bookshelf browsing searches. The infrequency of scans is another characteristic of bookshelf browsing searches.

F. REFINEMENTS IN SUBJECT SEARCHES

Refinement is defined as the application of a criterion to the output of a subject catalog search. Refinements occur either prior to the searcher's access of the catalog or during the course of the subject catalog search. Those occurring before the search are termed a priori refinements and those occurring in the course of the search are termed interactive refinements. The type and number of refinements recorded in 189 subject searches are enumerated in Table 23 and ranked in order of their occurrence in subject searches. In terms of frequency of occurrence per search, a priori refinements were more plentiful and were recorded in three of every four subject searches, whereas interactive refinements were recorded in approximately one of every three subject searches.

A priori refinements represent criteria which the searcher applies when determining the relevance of a specific catalog item and which are invoked before the searcher accesses the catalog. Examples of a priori refinements are given in the partially coded protocol of Figure 15. The searcher's expressed topic is "Hot Air Balloons" and the initial access point is "Balloons." The searcher implies that he is particularly interested in books on pre-1945 airships. Should the searcher scan the catalog and find an item whose textual information contains the term "airships" and/or a reference to

```
42 Results When Looking For Power Resources

1. Title:           The poverty of power: energy and the economic
                    crisis
   Other Subjects:  Power (Mechanics)
                    Energy policy -- United States
   Call Number:     HD 950 A2 C643 1976

2. Title:           Beyond the age of waste: a report to the Club
                    of Rome
   Other Subjects:  Economic History -- United States
                    Natural Resources
                    Energy Policy -- United States
   Call Number:     HC 59 G231 1981

   Print More?
```

Figure 13. Example of an Intermediate Display of a Retrieved Item in an OPAC Subject Search

DC: OK. What are you looking at on the card?
1A S: "Trance."
DC: OK. You are looking at the top?
S: Yeah.
DC: OK.
S: And I'm not finding it.
3B DC: OK. Are you still looking at the top of the card?
S: Yeah. Well, I'm going to go over to the religious section ....
DC: OK. Now what are you looking at on the catalog?
1C S: I'm looking at the headings, looking at "Religion, Primitive." I'm
3B not really sure what I'm gonna find.
DC: You are looking at the titles?
S: Yeah.
DC: Just read it out loud what it is that you are flipping past there.
3C S: I'm not flipping past really anything. Check under "Religion, History."
DC: OK. Why do you think that you might find it there?
S: I'm looking for a particular type of religion, the Holy Ghost People and their trances, so I'm just checking to see if maybe they have something there.
DC: Where are you looking on the cards--at the title?
S: Yeah. I'm looking to see if maybe they mention something in their description.
DC: OK. Now you have gone to a different drawer.
S: Right. Just checking religions in general .... "Cults." Oh, this looks good. "Modern Religious Cults and Movements" might have something that
3C I want.
DC: Are you going to write down the book or the card number?
S: The book and the number, call number I guess.... This might also have something in dealing with trances. I'll take it down.
DC: OK. So you are continuing on now--looking for some more books?
S: I'm looking at the description again. (DC's Note: Points to the title)
DC: OK.
3C S: OK. This might have something. I'm not really sure what ....
DC: OK. And you're gonna write down the call number again.
S: Yes.
DC: OK. You're still looking through some more?
S: Doing the same thing ... Actually it's sort of changed over to different topics so I don't need any of this. Ah hah! (DC's Note: Lots of silence) Now, I'm checking down here because I'm not really sure what they are dealing with.
DC: OK. What are you reading--"Liberty, Authority, Democracy"?
S: Yeah. The stuff with the numbers down at the bottom.
3F DC: I don't know what they call them.
S: I don't know either. OK. "Seduction of the Spirit." Hmm--no.
DC: You are flipping through pretty fast. Are you just looking at titles?
3C S: I'm just seeing if maybe there is something that was just really general ....
I think what I'll do is go over and see if there is anything under the "Holy Ghost People."
DC: All right. Off to "Holy Ghost."
1E S: Holy Ghost.

Figure 14. Protocol of a Subject Search for Trance Dances of the Holy Ghost People Coded with Scans and Access Points

DC: I think the H's are somewhere along in here. There they are in the back. OK. Let's see.
S: "Holly"... "Holy See." Well, we'll check out the Holy Spirit.
DC: OK. Now are you looking at the descriptions again underneath the titles or ...?
S: Yeah. I'm not finding much.
3B DC: All right.
S: Where did I see that, it was talking about the "Holy Spirit in Christian Theology." Um. No, I don't think so.
DC: OK. Why don't you think that would fit?
S: Well, I don't know.
DC: Just doesn't sound right?
S: Just doesn't hit it. Actually I don't think any of this is really what I want. Now I'm just looking back at the top again.
3C DC: I think you're out of the Spirits.
S: Let me see where "Holy" starts. "Alliances." Maybe they've got a Holy Ghost here. Gee, "Holy Ghosts."
3B DC: Looking back again at "Holy Spirit"?
S: Oh. That's what I was under ... I'm not really sure that that's what I want though. I don't think that's what I want. I think what I'll do is check out these to see what they tell me.

Key:

DC = Data Collector
S = Searcher
1A = Access Point, partial representation of expressed topic
3B = Scan, subject headings
1C = Access Point, different from expressed topic
3B = Scan, subject headings
3C = Scan, titles
3F = Scan, t[illegible]ngs
3C = Scan, titles
1E = Access Point, partial representation of expressed topic, not found
3B = Scan, subject headings
3C = Scan, titles
3B = Scan, subject headings

Figure 14.--Continued

Table 22. Chi-Square Test of the Relationship between Number of Scans and Access Points

Null Hypothesis: Number of scans and access points are independent.
Alternate Hypothesis: Number of scans is dependent on the number of access points.
N = 189, 2x3 Table, .01 Level, Degrees of Freedom = 2, Chi-Square $\geq$ 9.21:R

Frequency Row Percent Column Percent	Number of Scans			
	0 or 1	2	3 or more	
Number of Access Points = 1	62 50% 84%	24 19% 65%	38 31% 49%	124
Number of Access Points $\geq$ 2	12 18% 16%	13 20% 35%	40 62% 51%	65
TOTAL	74	37	78	189

Chi-Square = 20.698*
*Rejects the null hypothesis at .01, 2 degrees of freedom

Thus, there is a significant relationship between the number of access points and scans.

Table 23. Types of Refinements Ranked by Frequency of Occurrence

Type of a priori Refinement	Frequency	Percentage
2A Keyword	111	28.8
2B Key phrase	109	28.2
2E Publication Date	36	9.3
2D Key name and/or title	29	7.5
2I Negative keyword or phrase	21	5.4
2F Language	13	3.4
2K Dates covered	12	3.1
2J Illustrations	9	2.3
2C Key concept	9	2.3
2G Library location	8	2.1
2P Level of specificity	8	2.1
2L Geographic areas	7	1.8
2R Knowledge level	7	1.8
2N Have something similar	2	0.5
2O Number of items needed	2	0.5
2H Number of pages	1	0.3
2S Time factor	1	0.3
2T Other	1	0.3
2Q Unknown	0	0.0
2M Format	0	0.0
TOTAL	386	100.0%

Type of Interactive Refinement	Frequency	Percentage
4I Negative keyword or phrase	45	30.6
4B Key phrase	42	28.5
4A Keyword	31	21.1
4N Have something similar	8	5.4
4P Level of specif_city	6	4.0
4D Key name and/or title	5	3.4
4C Key concept	2	1.4
4J Illustrations	2	1.4
4M Format	2	1.4
4E Publication date	1	0.7
4G Library location	1	0.7
4Q Unknown	1	0.7
4R Knowledge level	1	0.7
4F Language	0	0.0
4H Number of pages	0	0.0
4K Dates covered	0	0.0
4L Geographic areas	0	0.0
4O Number of items needed	0	0.0
4S Time factor	0	0.0
4T Other	0	0.0
TOTAL	147	100.0%

DC: Where are you going to start?
S: I'm looking for balloons.
DC: Where on the card are you looking?
S: Looking at the top left hand corner. Well, I'll first find "Balloon" but then I'll look at the top left hand corner for the number ... for wherever it is at.
1A
DC: OK. Are you looking for a particular book? Do you have a call number in mind?
S: No. I'm looking for balloons and then I'm gonna ... look. See, I'm looking at it in reference to airships and stuff like that. It's for a term paper.
2A
DC: All right. OK, you're flipping past a few cards. Where are you looking now?
S: I'm looking at the top.
DC: Still at the top? OK.
S: I'm avoiding ... because I have to have stuff before 1945.
2K
DC: OK. So date's important then.
S: Little bit ....

Key:

DC = Data Collector
S = Searcher
1A = Access Point, partial representation of expressed topic
2A = A priori Refinement, keyword (i.e., "airships")
2K = A priori Refinement, dates covered (i.e., "before 1945")

Figure 15. Coded Extract of a Protocol of a Subject Search for Airships with a priori Refinements

"before 1945" or "pre-World War II" in the title, notes, or subject headings, the searcher would presumably judge that item to be relevant.

Interactive refinements also represent criteria which the searcher applies when determining the relevance of an item. They are invoked by the searcher during his access of the catalog, especially when scanning catalog information. The searcher whose protocol is displayed in Figure 16 consulted the catalog for her expressed topic, "Prenatal Care." In the course of the search, she indicates she is particularly interested in books on the health aspects of prenatal care that have been recently published and are written in English. Such details emerge both in the beginning and middle of her search, but they appear to be criteria that were determined before her consultation of the subject catalog. In contrast, the fact that she is not concerned about the "husband's psychology" or "underdeveloped countries" and prenatal care emerges when she reads specific titles which cover those topics.

We have already seen that the majority of searchers altered their expressed topic when accessing the catalog by consulting an access point that had significantly fewer terms than the expressed topic, or was a partial representation of the expressed topic, or was different from the expressed topic. Since a priori and interactive refinements primarily deal with a specific term or phrase in which searchers are interested (see Table 23), searchers may be using refinements as extensions of the access point or search argument. Simply put, refinements provide the subject catalog searcher with a means of including additional terms in the access point. This approach simulates the "stringsearch" of SDC's ORBIT retrieval system and the Boolean search capability of Claremont College's OPAC. In both instances, (explicitly and implicitly, respectively) the searcher creates a partition of the database containing a particular phrase or term. This partition is then searched for the occurrence of another phrase or term.

In both protocols featuring refinements (Figures 15 and 16), the catalog searcher initially partitions the file (e.g., balloons, prenatal care), then scans the partition for a keyword or phrase (e.g., airships, health). Library catalog searchers are thereby implicitly setting up a Boolean "AND" operation. The frequency of negative interactive refinements is evidence that subject searchers are simulating the Boolean "NOT" operation in the course of their search, especially after having perused the subject content of items found under the access point. The searcher's remark that he wanted books about "airships and stuff like that" (Figure 15) implies that he will be interested in catalog items containing the words "blimps," "dirigibles," and "lighter-than-air aircraft;" the Boolean "OR" operator is implicit in the searcher's remark.

Searchers who employ a priori and interactive refinements are often extending the vocabulary of the initial access point; consequently, the initial access points in searches containing refinements are likely to be those in which searchers have consulted a partial or different representation of their expressed topic. Searchers whose initial access points are the same as the searchers' expressed topics can be expected to contain few refinements because searchers have successfully matched the vocabulary of their expressed topics with the catalog's controlled vocabulary. In Table 24 a significant relationship between initial access point and number of a priori refinements is reported. The percentage of searches in which the initial access point was

S: I'm looking for prenatal care. Should be in here. Here we go. Prenatal Influences, Prenatal Care. OK, at this point I would go upstairs and look up the book.

DC: OK. Are you looking for a particular book or are you ...?

S: Just basic prenatal care. And in English .... I would be more interested in something like this, as opposed to a history or a major study. I'm looking this up for myself ....

2F
2I

DC: So you are writing down the call number or the title or author?

S: Just the author. I don't know how it would work in the Home Ec library.

DC: I see.

S: All right. I think this is all the subjects so ... so this book is here. That would be good.

DC: "Birth facts and legends." History-oriented.

S: Yeah.

DC: "Life before birth." What makes you think that that might be useful?

S: Just because I might be pregnant now and I just want to know what's happening.

DC: OK.

S: This one is real old. 1905 is a pretty ancient book.

DC: Indeed. You seem to be reading the card in some detail.

2E

S: Here's one in French so that won't help. And this one--"What every husband should know about having a baby"--I'm not too worried about that right now. I don't feel that I need to read up about his psychology. I'm just worried about the health aspects at this point. So I'm mostly looking for ways to take care of myself toward that .... This one I'm not sure if I am interested in because it really doesn't give me much information on the card.

4F
4I
2A

DC: And it is a 1930 book.

S: This one was rewritten in the '50s.

DC: Yeah.

S: I'll scan further down. Probably I'll stop about .... I'll just breeze through and see what I can find from these. This might not be too bad --"Your baby is on the way."

DC: That's 1961.

S: This is pretty old--1924. See this one? The only thing I wonder is whether it has to do with underdeveloped countries.

4I

DC: It's by the World Health Organization.

S: Symposiums .... Well, I guess I'm going to stop here.

Key:

DC = Data Collector
S = Searcher
2F = A priori Refinement, language (i.e., "English")
2I = A priori Refinement, negative key (i.e., "history")
2E = A priori Refinement, publication date
4F = Interactive Refinement, language (i.e., "not in French")
4I = Interactive Refinement, negative key (i.e., "husband's psychology")
2A = A priori Refinement, keyword (i.e., "health")
4I = Interactive Refinement, negative key (i.e., not in "underdeveloped countries")

Figure 16. Coded Protocol of a Subject Search for Prenatal Care with a priori and Interactive Refinements

Table 24. Chi-Square Test of the Relationship between Type of Initial Access Point and Number of a priori Refinements

Null Hypothesis: Type of initial access point and number of a priori refinements are independent.
Alternate Hypothesis: Type of initial access point and number of a priori refinements are dependent.
N = 148, 3x3 Table, .01 Level, Degrees of Freedom = 4, Chi-Square $\geq$ 13.277:R

Frequency Row Percent Column Percent	Number of a priori refinements per search: 0	1 - 2	3 or more	
Type of Initial Access Point:				
1A Partial representation of expressed topic	11 15% 28%	34 47% 52%	27 38% 61%	72
1B Same as expressed topic	26 48% 67%	19 35% 29%	9 17% 21%	54
1C Different from expressed topic	2 9% 5%	12 55% 19%	8 36% 18%	22
TOTAL	39	65	44	148

Chi-Square = 22.14*
*Rejects the null hypothesis at .01, 4 degrees of freedom.

Thus, the number of priori refinements and type of initial access point are dependent.

a partial or different representation of the expressed topic and in which there were no a priori refinements contrasts with the percentage of searches in which the initial access point was the same as the expressed topic and in which there were no a priori refinements. The greater the number of a priori refinements, the less likely that the initial access point will be the same as the expressed topic. The chi-square test was also used to find a relationship between initial access point and interactive refinements and a significant relationship was noted (Table 25).

Most likely, the occurrence of a priori or interactive refinements in subject searches, particularly those which involve a term or phrase, namely a keyword, key phrase, and/or negative keyword or phrase, represent searchers' makeshift methods of postcoordinating the terms used to represent the subject of their expressed topic. The earlier discussion of access points pointed out that subject searchers have "learned" to simplify their expressed topic by reducing the number of terms in the access point and, occasionally, employing a term or phrase which is conceptually broader or more general than the expressed topic. While accessing the subject catalog under the general term, searchers are keeping in mind other terminology which, if it appears in the catalog during the search, will mark a potentially relevant item.

This crude attempt to postcoordinate terminology in subject searches has not been observed in previous catalog use studies but bears similarities to a process that previous catalog studies have dubbed "browsing." Researchers have characterized the subject search as a dynamic interaction in which the "'information need' tends to be modified, to a greater or lesser extent, by what is found during the search, and the final set of documents, accepted by the searcher as 'useful' in relation to his requirements, may be somewhat different from the 'kinds' of documents he visualized as useful when the search commenced." (Lancaster, 1968, p. 123.) What Lancaster has described is the end result of a subject search in which "browsing" has been the principal activity.

> Persons browse, among other reasons, because their criteria of interest are not precisely defined, or are open-ended and subject to redefinition. They may desire to locate and actively seek out previously undiscovered sources which satisfy those criteria, whatever the publication format or "discipline of residence" of these sources turns out to be. Taken together, these factors largely constitute what we characterize as "the spirit of inquiry." (Hildreth, 1982c, p. 185.)

In Figure 17, a coded protocol of a subject search depicts the browsing process that Hildreth describes. The searcher's spoken thoughts disclose the open-endedness of her expressed topic. In this example, the searcher uses refinements to qualify the search so that the output of the search (i.e., scanned items) does not veer towards nonrelevant items.

The use of refinements to assist in browsing, in the manner of the searcher in Figure 17 or like Lancaster's description, was infrequent in this study. In contrast, refinements were usually attempts to introduce a postcoordinated subject term or phrase into the course of the search. We can only speculate that searchers' limited success in matching the vocabulary of their expressed topics with that of the subject catalog had conditioned them to consult brief

Table 25. Chi-Square Test of the Relationship between Type of Initial Access Point and Number of Interactive Refinements

Null Hypothesis: Type of access point and number of interactive refinements are independent.
Alternate Hypothesis: Type of access point and number of interactive refinements are dependent.
N = 148, 3x3 Table, .01 Level, Degrees of Freedom = 4, Chi-Square $\geq$ 13.277:R

Frequency Row Percent Column Percent	Number of Interactive Refinements: 0	1	2 or more	
Type of Initial Access Point:				
1A Partial representation of expressed topic	36 50% 40%	13 18% 50%	23 32% 74%	72
1B Same as expressed topic	42 78% 46%	6 11% 23%	6 11% 19%	54
1C Different from expressed topic	13 59% 14%	7 32% 27%	2 9% 7%	22
TOTAL	91	26	31	148

Chi-Square = 15.876*
*Rejects the null hypothesis at .01, 4 degrees of freedom

Thus, the number of interactive refinements and type of initial access point are dependent.

S: I have no idea where I'm going to start. I didn't really come with a purpose in mind. I'm just looking maybe for something.
DC: OK, go ahead. What type of general thing are you looking for?
S: Small Business. So I don't know.
DC: We are not so much interested in what you are looking for as much as how you go about looking for it. As you flip through, you are looking
1A at the titles?
S: No, at what the books contain.
DC: I see.
S: Such as commerce or industry or ....
DC: Is there something particular about that card? You seem to be holding it out, "The lady means business."
S: No particular reason, I'm interested in the points of women being in business.
2B DC: So you are going to write down the call number then?
S: Yeah, but it doesn't go on from here to tell you what the book is about; it just tells you the author.
DC: Right. That one is of interest because it deals at least with small business? OK, (you're) flipping through some more cards ... "Introduction to business." Why not that one?
S: Because it says it is a textbook for the first course in business on a
4I collegiate level and I've already had that type of course ... "Selecting and operating a business of your own: A guide for choosing and setting up a small business of your own."
DC: Why is that not of interest?
S: Because it says 1896.
DC: Not real current.
2E S: Yes.
DC: You've passed up a couple on starting your own business. You're not too interested in that angle?
S: Well, this one is about doing it for $1,000 or less, so I'm certain it is just a spare time fortune guide. I guess there is a certain
4B business aspect that I am looking for--"Business occupations in retail trade." I don't know what that would be. I have a list so I'll go and look at some of them.
DC: Still looking up at the top and sort of down beneath?
S: Yes.
DC: OK. (You're) looking through some more cards. That says "Small Business" right in the title ... looking through more now ... "Business Adventures."
S: I think I am going to stop soon so I can find a book .... "Business Law," "Business Loans." I was looking possibly for a book on small business loans. Also grants.
DC: OK.
S: That is the only book I see on small business loans. I'm going to look
2B now.

Key:

S = Searcher
DC = Data Collector
1A = Initial Access Point, partial representation of expressed topic (i.e., "Business")
2B = A priori Refinement, key phrase (i.e., "Women in Business")
4I = Interactive Refinement, negative term or phrase (i.e., "Textbook for First Course in Business")
2E = A priori Refinement, publication date
4B = Interactive Refinement, key phrase (i.e., "spare time fortune guide")
2B = A priori Refinement, key phrase (i.e., "Small Business Loans or Grants")

Figure 17. Coded Protocol of a Subject Search for Small Business with a priori and Interactive Refinements

versions or variations of their expressed topic in the catalog and to screen catalog items for the occurrence of a particularly relevant term or phrase.

Such subject catalog searchers' makeshift postcoordination indicates a need to build in automatic postcoordination as a feature of OPACs. Minimally, OPACs should allow searchers to alter the results of a single search argument by a Boolean or Boolean-like operation. Such a feature could be incorporated subtly so that searchers are not overwhelmed by what spokesmen (Summit, 1980; Garfield, 1980; Kugel, 1971) have considered a complex operation. It remains for research on activity reports and transaction logs of OPACs that feature explicit Boolean searching to verify what has merely been speculation about the complex nature of Boolean operations.* Transaction logs differ from the analysis of protocols in this study, as they are a machine-readable record of the interaction of searcher and system and provide a more accurate record of the online search than the analysis of protocols because of the subjectivity of assigning constituent elements to protocols. Constituent elements are already "assigned" in transaction logs through the searcher's "logicalizing" of his thoughts and physical actions into messages understandable to searcher and system alike.

Because of the frequency of a priori and interactive refinements which are associated with searching vocabulary, other types of refinements seem less important in subject catalog searches. Refinements such as publication date, level of specificity, dates covered, format, and languages are familiar to searchers of databases available through commercial retrieval systems as "limits." We expect that such refinements or limits will become more frequently incorporated into subject searches when postcoordination is allowed by OPACs. Since the searcher's principal objective is to find items relevant to a subject and employ every means possible to manipulate searching vocabulary, criteria such as format, language, and publication date play a secondary role.

G. COMPLETING THE SUBJECT SEARCH: AIDS, FLAGS, RECORDS, AND EXITS

Few subject searchers obtained assistance from the librarian during their catalog search. This may be attributed to the data collection method itself; searchers might have been reluctant to approach the librarian because one additional individual was involved already in what is usually a solitary activity. Also, searchers might have considered such a request an admission of failure. Of the fourteen searches in which aid from the librarian was requested, three searches involved multiple requests for assistance.

Even though aid was infrequently sought, those searches in which it was sought contained a significantly greater number of access points than those searches in which it was not (see Table 16). Instances in which aid was requested are enumerated in Table 26, accompanied by a description of how the librarian helped the searcher.

---

*Currently under investigation by Charles R. Hildreth of OCLC and Edwin Brownrigg of the University of California through grants awarded by the Council on Library Resources.

Table 26. Frequency and Type of Aid Requested from the Librarian

No. of Aids	Expressed Topics	Type of Aid Given
6	Life after death Dreams Commerce and geography of Poland Draft: conscription of the military Mountain men	Suggest subject headings
5	Mountain men Infant and maternal health in China, India, and U.S. Women, Infant, and Children Agency Twins--bonding, psychological effects of separation	Unknown
2	Racquetball/Bowling Alice Aycock	Give spelling
2	Pictures of the 1940s Geology of the Adirondacks	Suggest best sources
1	Rules and regulations for small business	Provide reference book
1	Geography and commerce of Poland	Tell how to use card catalog

Requests for assistance were few, and the most frequent requests dealt with vocabulary for subject searching. Online displays of the catalog's controlled vocabulary are suggested as features for OPACs to help searchers who cannot find just the right term or phrase or are having difficulty matching the controlled vocabulary with the vocabulary of their access point(s) (see Chapter IV, Section E). Online vocabulary displays featuring the keyword-in-context concept can help searchers whose vocabulary is embedded in a subject heading and does not match with the first word of a subject heading. The ability to retrieve alphabetical and keyword-in-context displays in response to a truncated term or phrase could also assist OPAC searchers who are poor spellers.

Flagging (the physical or mental marking of a specific location in the catalog) occurred infrequently (15 times), but the fact that it occurred implies that catalog searchers sometimes maintain a physical or mental record of where and what they have searched in the subject catalog. The subject catalog provides the searcher with physical clues about what has been searched and where (e.g., guide cards, drawers, alphabetical ranges, and individual catalog records); however, the OPAC, particularly one using a CRT, will not give the subject searcher as many physical clues as the library catalog unless hard copy is recorded simultaneously and is available within arm's reach or the OPAC itself maintains a record of terms and phrases which have been searched. A feature similar to DIALOG's "display sets" command or ORBIT's "history" command is recommended for OPACs so that searchers can periodically review what they have done. Likewise, online displays of the controlled vocabulary (e.g., alphabetical, keyword-in-context, or related term displays) must allow searchers to move forward and backward in the list, just as subject catalog searchers can presently go forward and backward by flipping cards in the catalog or by manually turning a knob on a COM reader. Searchers can also estimate the number of items assigned a particular heading by physically marking in some way a block of catalog cards sharing the heading. To equal such a capability in OPACs calls for postings information to be incorporated into online vocabulary displays and included in response to a search command or search operation.

In Table 27, information which searchers recorded at the catalog is summarized. There are no surprises. In nearly two-thirds of the searches, the call number or the call number area(s) were recorded or memorized. Other frequent elements recorded by searchers were title and author. There were only 13 instances when searchers recorded full bibliographic information. Such findings indicate how few data elements are needed (or wanted) in online bibliographic displays of individual items. More importantly, the figures for call numbers, particularly memorized call numbers, disclose how frequently bookshelf browsing searches occur.

Of the 189 subject searches, six ended in a request for assistance from the librarian. In the remaining searches, the searcher exited to the stacks (78.7%), from the library (8.2%), to study tables (1.6%), or to another library (1.1%). Thirteen searchers (7.1%) consulted other indexes or library sources. Their searches dealt with a current event (e.g., Boat People) or a reference request (define a medical term (Atleio), locate a criticism of a book, etc.).

Table 27. Frequency of Bibliographic Information Recorded by 189 Subject Searchers

Type of Record	Frequency	Percentage
7B Call number	83	33.3
7H Unknown	65	26.1
7G Memorized call number or call number areas	34	13.7
7F Author, title, and call number	14	5.6
7I Full bibliographic information	13	5.2
7D Title and call number	12	4.8
7E Title and author	12	4.8
7J Author and call number	11	4.5
7A Title	3	1.2
7C Author	2	0.8
TOTAL	249	100.0%

H. CONCLUSION

Data analysis of subject searchers' protocols resulted in a description of participating libraries and the characteristics of their searchers. Data collection at one university, three public, four high school, and four college libraries yielded 228 protocols; 39 protocols were eliminated from the study because the data collector may have biased the protocol by influencing the search strategy.

No surprises emerged in the analysis of subject catalog searchers' educational backgrounds and demographic characteristics. The mean age of searchers in public libraries was greater than that in high school, college, and university libraries. Public library searchers also displayed more variety in terms of the highest grade of schooling completed than searchers at other libraries. High school, college, and university subject catalog searchers visited the library more often than public library searchers, which is probably a reflection of the daily routine of coursework requirements. Subject catalog searchers at all four types of libraries were likely to have had training in library catalog searching.

The analysis of constituent elements in coded subject search protocols disclosed that over two-thirds of the searches contained less than the average number of elements. Searches featuring fewer than the average number of constituent elements were indicative of bookshelf browsing searches when this characteristic was paired with one access point and/or scan and/or the recording or memorization of a call number or call number area. The predominance of bookshelf browsing searches suggests features for OPACs which incorporate shelf browsing into the actual online search through an online tool comparable to the printed Subject Guide to the Collection of the Public Library of Columbus and Franklin County or displays of the classification numbers and their corresponding captions. Because of the frequency of bookshelf browsing searches in public, high school, and college libraries, OPAC enhancements such as the Subject Guide are recommended for these types of libraries.

The analysis of access points verified the findings of previous catalog use studies concerning searchers' perseverance. On the average, one of every three searchers consulted more than one access point, but those who continued their search had probably not found their access point in the catalog on the initial attempt. Subsequent attempts were marked by searchers' consultation of the catalog with terminology that was different from or a partial representation of their expressed topic. Few searchers sought the exact representation of their expressed topic as a subsequent access point when accessing the catalog after their initial attempt, even when their initial access point was not an exact representation of their topic.

A comparison of the vocabulary of access points with expressed topics showed that vocabulary frequently changed between the time the searchers approached the catalog and the time they accessed it. Specifically, this change was identified with a significant reduction in the number of terms used to represent the access points. The change was also marked by increased generality of the vocabulary of access points, particularly when it constituted a partial representation of the expressed topic (e.g., access

point=Sports, expressed topic=Russian Sports). One possible explanation for this change from expressed topics to access points is that, in previous subject searches when using the exact representation of expressed topics, searchers failed because the search vocabulary was too specific or detailed.

The fact that high school students, the youngest searchers in this study, were prone to access the catalog with the exact representation of their expressed topics helps to support the explanation; such expressed topics were simple, general, and brief (i.e., similar in form to access points which had undergone the "change" process). Perhaps as library catalog searchers advance in the course of their educational careers their expressed topics may become more specific and complex in terms and facets, but they may continue to rely on the simple and brief statements that were so successful in their early subject catalog searches and translate their expressed topics into such statements.

Catalog searchers in this study infrequently sought assistance from the librarian. The significant relationship between the number of access points and requests for librarian assistance in subject searches, and a ranking of the type of assistance requested, demonstrates the need for online displays of the catalog's vocabulary in OPACs to assist searchers in the selection of subject terminology (e.g., alphabetical, keyword-in-context, and/or related term displays).

Searchers' scanning of subject information in the catalog is an indication of the data elements necessary in intermediate OPAC displays of bibliographic items (e.g., call number, title, subject headings, and notes). Few searchers scan or record the full bibliographic information of library material in a subject search of the library catalog. The predominance of bookshelf browsing searches and recording of call numbers or call number areas strongly support the inclusion of call numbers in all displays of bibliographic information of library material.

Since a priori and interactive refinements primarily dealt with a specific term or phrase, searchers' use of refinements centered on extending the vocabulary of the access point or search argument. Their approach is comparable to the "stringsearch" of the ORBIT retrieval system or the Boolean search capability of Claremont College's OPAC. A significant relationship between refinements in subject searches and type of initial access point supports the notion that refinements are searchers' makeshift attempts to postcoordinate terminology in subject searches of the library catalog. The approach was compared with explanations of browsing in previous catalog use studies; few searchers in this study employed browsing in the traditional sense. In contrast, they seemed to be occupied with perusing catalog information for the occurrence of a term or phrase that would indicate an item's relevance or usefulness. So occupied were subject searchers that typical refinements or "limits" (e.g., date of publication, language, and format) were infrequent. Subject catalog searchers' makeshift postcoordination indicates a need to build in automatic postcoordination as a feature of OPACs. Minimally, OPACs should allow searchers to alter the results of a single search argument by an implicit or explicit Boolean operation.

The analysis of frequency of search activity codes and their constituent elements resulted in an understanding of how library catalog searchers perform

subject searches. We have also gleaned from searchers' spoken thoughts and mental or physical actions features for OPACs which must be present to support what searchers presently do at the library catalog. The following section describes the construction of a post hoc process model of subject searching based on the analysis of search protocols. Additional features from this model are recommended for OPACs to support the present search tactics of library catalog searchers.

IV. DISCUSSION: AD HOC PROCESS MODELS OF SUBJECT SEARCHING AND IMPLICATIONS FOR POWER

A. INTRODUCTION: AD HOC PROCESS MODEL OF SUBJECT SEARCHING

An ad hoc process model of subject searching was constructed from the analysis of 189 subject search protocols (Figure 18). This model serves as a composite representation of the process of subject searching the library catalog. It was constructed by examining the sequence of constituent elements assigned to every subject search. If the sequence of constituent elements of a subject search did not fit the model, the model was altered to accommodate the search and all others that had been examined. The ad hoc model represents the process of subject searching in a flowchart manner. No one subject search fits the composite model exactly, however, every subject search matches a portion of it.

Portions of the composite model which occurred frequently are depicted in subsets of the model. These subsets are more faithful representations of the process of subject searching than the composite model. The latter is indicative of all subject searches, but it does not portray any one search in particular. Subsets of the composite model disclose the variety of approaches to searching the library catalog. Subsets of the ad hoc process composite model are identified in the following sections. They are illustrated by a subject search protocol, depicted in a flowchart manner, and discussed in view of how they presently support the search tactics of subject searchers. Such search tactics are features required of online public access catalogs to support the present search tactics of subject searchers and the variety of approaches to subject searching employed by searchers of the library catalog.

The OPAC has the potential to support features never possible in subject searching of the library catalog. The analysis of search protocols and the construction of ad hoc process models brought to light features for OPACs which are not possible in searches of the library catalog. For example, the discussion of refinements in subject searches (see Chapter III, Section F) disclosed searchers' makeshift attempts to perform Boolean operations at the library catalog. We will see how searchers' failures to match their subject terminology with that of the catalog, searchers' use of the catalog to find class number areas for browsing at the bookshelf, and other incidents in, or frequent processes of, subject searching reveal features for implementation in OPACs that were never possible at the library catalog. Features that support the present search tactics of library catalog searchers and those that are now possible, thanks to technology, add to the online public access catalog the potential to match or surpass the capabilities of the traditional library catalog. The user-oriented interface is the crucial communications component of OPACs, through which the searcher unleashes the system's power.

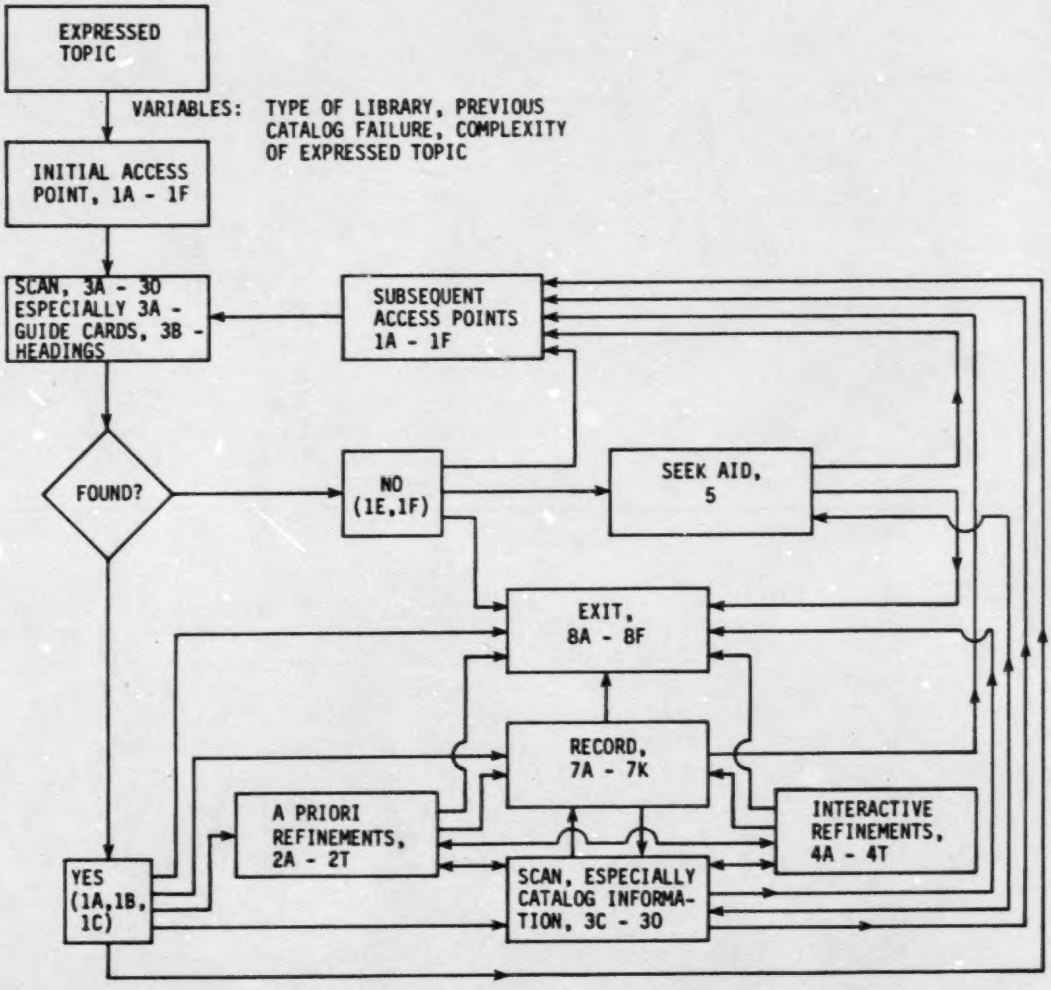

EXPRESSED TOPIC
VARIABLES: TYPE OF LIBRARY, PREVIOUS CATALOG FAILURE, COMPLEXITY OF EXPRESSED TOPIC
INITIAL ACCESS POINT, 1A - 1F
SCAN, 3A - 3O ESPECIALLY 3A - GUIDE CARDS, 3B - HEADINGS
SUBSEQUENT ACCESS POINTS 1A - 1F
FOUND?
NO (1E,1F)
SEEK AID, 5
EXIT, 8A - 8F
RECORD, 7A - 7K
A PRIORI REFINEMENTS, 2A - 2T
INTERACTIVE REFINEMENTS, 4A - 4T
YES (1A,1B, 1C)
SCAN, ESPECIALLY CATALOG INFORMATION, 3C - 3O

Figure 18. Ad Hoc Process Model--Composite Model

B. BOOKSHELF BROWSING SUBJECT SEARCHES

1. Description.

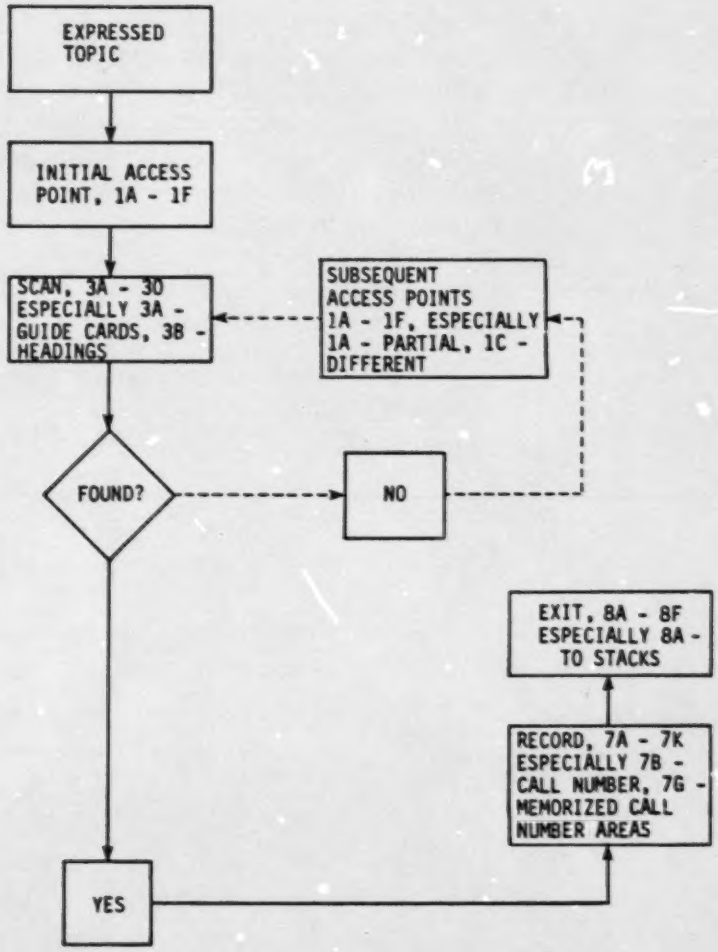
The searcher's objective in bookshelf browsing searches is to link his subject with at least one area of the classification scheme which is particularly prolific in terms of both the number of books assigned to the class number and those matching the searcher's topic. Figure 5 is an example of such a search--one in which book selection occurs at the bookshelves, not at the library catalog. Outstanding characteristics of bookshelf browsing searches are: a single access point, fewer than average (9.73) constituent elements, and a record element for a written (7B) or memorized (7G) call number or call number area. Occasionally, when the searcher's initial access point does not match the controlled vocabulary of the subject card catalog, the searcher immediately selects another access point which is a partial (1A or 1E) or different (1C) representation of the expressed topic (Figure 19).

2. Required Features.

To support the bookshelf browsing approach of subject searchers, online public access catalogs must allow searchers to enter queries in the vocabulary of their initial access point. In response, the OPAC produces a list of the subject catalog's controlled vocabulary terms and phrases which, like terms in the subject catalog, are alphabetically adjacent or nearby the input terminology. This alphabetical display should contain subject headings, number of postings per subject heading, and the class numbers in common to the majority of the books assigned a particular subject heading. Just in case the searcher's input term or phrase may not be listed in the alphabetical display, the OPAC must tell the searcher how to manipulate the display so that he can reverse or advance in the alphabetical display. The Ohio State University's Library Control System (LCS) produces a subject heading display which, except for the omission of class numbers common to the majority of books assigned a particular number, resembles what subject searchers scan in bookshelf browsing searches at the library catalog. The LCS display is depicted in Figure 20; class number areas are appended to approximate what a bookshelf browsing display should look like. Implemented in the OPAC, the display consolidates the scan and record steps of bookshelf browsing in the subject library catalog into a single, simultaneous online display of subject headings and class numbers.

3. Beyond the Library Catalog: Bookshelf Browsing Features for OPACS.

The make-up of Figure 20 is somewhat misleading, since library catalogs do not tally the number of books assigned to a particular subject heading nor, obviously, do they display the class number assigned to the majority of titles assigned a particular subject heading. However, the searcher "gets a feel" for such information by scanning, manipulating, and comparing guide cards and ranges of cards prefaced by the same subject heading. Providing the number of postings and the predominant class number in bookshelf browsing displays would give OPAC searchers more information than presently available in subject catalog searches; thus, the OPAC would perform an operation that is impossible for the static and fixed catalog to accomplish every time a searcher consults and matches a particular subject heading.


EXPRESSED TOPIC
INITIAL ACCESS POINT, 1A - 1F
SCAN, 3A - 3D ESPECIALLY 3A - GUIDE CARDS, 3B - HEADINGS
SUBSEQUENT ACCESS POINTS 1A - 1F, ESPECIALLY 1A - PARTIAL, 1C - DIFFERENT
FOUND?
NO
EXIT, 8A - 8F ESPECIALLY 8A - TO STACKS
RECORD, 7A - 7K ESPECIALLY 7B - CALL NUMBER, 7G - MEMORIZED CALL NUMBER AREAS
YES

(Note: Dotted line indicates process which frequently occurs but did not occur in the search modeled.)

Figure 19. Ad Hoc Process Model of a Bookshelf Browsing Subject Search

User: SUB/ISLAM -- ADDRESSES, ESSAYS, LECTURES

TBL	NO. OF CLASS AREAS	HITS	SUBJECT
11	PL216	4	ISLAM IN LITERATURE
12	BP20	26	ISLAM--ADDRESSES, ESSAYS, LECTURES
13	BP64	7	ISLAM--AFRICA
14	DT4	1	ISLAM--AFRICA--ADDRESSES, ESSAYS, LECTURES
15	Z7835	1	ISLAM--AFRICA--BIBLIOGRAPHY
16	DS38	1	ISLAM--AFRICA--CONGRESSES
17	209	1	ISLAM--AFRICA, NORTH--HISTORY
18	Z7835	1	ISLAM--AFRICA, SUB-SAHARAN--BIBLIOGRAPHY
19	BP64	3	ISLAM--AFRICA, WEST
20	BP64	1	ISLAM--AFRICA, WEST--ADDRESSES, ESSAYS, LECTURES
	PAGE 2 OF 4		FOR OTHER PAGES, ENTER PS AND PAGE NUMBER. FOR TITLES ON A SPECIFIC SUBJECT ENTER TBL/ AND LINE NUMBER

Figure 20. Alphabetical Subject Heading Display with Class Numbers for OPAC Bookshelf Browsers

User: CLA/ISLAM -- ADDRESSES, ESSAYS, LECTURES

CBL	NO. OF CLASS AREAS	SUBJECT	CLASS AREAS & (NO. OF BOOKS)
11	4	ISLAM IN LITERATURE	PL216, PR149, PQ283, PQ6066
12	12	ISLAM--ADDRESSES, ESSAYS, LECTURES	BP20(5), 297(5), BP165(4), BP25(2), BP161(2), BP170(2), BM42, BP49.5, BP88, DS33, DS35.4, 910
13	4	ISLAM--AFRICA	BP64(4), BP62, 276, 297
14	1	ISLAM--AFRICA--ADDRESSES, ESSAYS, LECTURES	DT4
15	1	ISLAM--AFRICA--BIBLIOGRAPHY	Z7835
16	1	ISLAM--AFRICA--CONGRESSES	DS38
17	1	ISLAM--AFRICA, NORTH--HISTORY	209
18	1	ISLAM--AFRICA, SUB-SAHARAN--BIBLIOGRAPHY	Z7835
19	2	ISLAM--AFRICA, WEST	BP64(2), BP172
20	1	ISLAM--AFRICA, WEST--ADDRESSES, ESSAYS, LECTURES	BP64
	PAGE 2 OF 4	FOR OTHER PAGES, ENTER PC AND PAGE NUMBER. FOR TITLES ON A SPECIFIC HEADING ENTER CBL/ AND LINE NUMBER/AND NO. OF CLASS AREA, E.G., CBL/12/5.	

Figure 21. Suggested Online Display of the Subject Guide to the Collection of the Public Library of Columbus and Franklin County

The Subject Guide to the Collection of the Public Library of Columbus and Franklin County provides the number of postings assigned to Dewey Decimal classification numbers in an alphabetical compilation of subject headings and cross references (Figure 6). In an online mode, the Subject Guide would give more details about class number areas than the example given in Figure 20. It would enumerate all the class number areas assigned the same subject heading and the number of books classed in each area. This is depicted in Figure 21, an adaption of the online LCS display and Subject Guide entry under "Islam--Addresses, essays, lectures." The searcher would learn from the online display in Figure 21 that over 50% of the books assigned this subject heading are shelved in three class areas--BP20, 297, and BP165. Online displays for bookshelf browsers, such as the examples of both Figures 20 and 21, would direct searchers to the most fruitful class number areas through a single display. To accomplish a comparable task in the card catalog, searchers would have to tally on paper or in their memory the class number areas assigned a particular subject heading as they encountered the class number areas in the catalog. Figures 20 and 21 illustrate how an OPAC can use its computational powers to perform an operation presently done by subject searchers in an ad hoc fashion at the library catalog.

C. SUBJECT SEARCHES WITH A PRIORI REFINEMENTS

1. Description.

In several bookshelf browsing searches, searchers employed a priori refinements. Refinements associated with vocabulary were introduced to restrict the meaning to a specific topic when the exact terminology of the topic, or a partial representation of the expressed topic, had been consulted but not found in the catalog. Figure 22 is a protocol of a subject search for an Audi repair manual in which the searcher did not find "Audi" in his initial consultation of the catalog. In a subsequent access of the subject catalog, the searcher found the subject heading "Automobiles--Maintenance and repair," but because the subject heading consulted included American-made cars, the searcher was mindful to look for the terms "European" or "foreign cars" as he scanned individual book titles in the library. The process of subject searching for the Audi repair manual is depicted in Figure 23.

This search, having a seemingly simple expressed topic and many of the characteristics of a bookshelf browsing search, resulted in failure. In fact, the searcher was more hopeful of finding the item in a bookstore than in the library catalog. A check of Maggie's Place, the OPAC of Pikes Peak Library District, to find books on "Audi" was also unsuccessful; however, searching the system under the subject heading "Automobiles, Foreign" produced the display shown in Figure 24. Such a search might have led the searcher to browse the bookshelves around 629.28.

In the protocol of the search for an Audi repair manual (Figure 22), the searcher uttered the phrase "European or foreign," a Boolean-like search argument. Had he been able to incorporate that phrase into the vocabulary of his access point, he might have retrieved items like the two manuals identified in the search of Maggie's Place. Instead, the searcher

S: What I am looking for is auto repair for an Audi automobile. Looking to see if there is anything under Audi in here, but I doubt it.
DC: You are flipping past the cards fairly quickly. Where are you looking?
S: A - U - D - I
DC: So you are focusing on the top part of the card?
1E S: Right ... nothing here. Now I'm going to Automobiles ... Automobile Engineering ... Automobile Repair ... Automobile Maintenance.
DC: Looking at the top part of the card?
1C S: Right. Automobile Maintenance. I doubt if I'm going to find it here.
DC: OK, where are you looking on the card now?
2B S: Well, something that says European or foreign cars or something that might give me ... there's a book on repairs of Audi that I've seen in bookstores and I'm just wondering if it's here.
3B DC: "Foreign Markets," "Survival Manual." Still looking at titles?
3C S: "Motor Maintenance," "Popular Mechanics." "Auto Repair for Dummies" -- that's good. "Establishing an Auto Repair Shop" ... so I may just give up.
8B

Key:

S = Searcher
DC = Data Collector
1E = Access Point, partial representation of expressed topic--not found (i.e., "Audi")
1C = Access Point, different from expressed topic (i.e., "Automobiles--Maintenance")
2B = A priori Refinement, key phrase (i.e., foreign cars)
3B = Scan, subject headings
3C = Scan, titles
8B = Exit, from library

Figure 22. Coded Protocol of a Subject Search for Audi Repair Manual with a priori Refinements

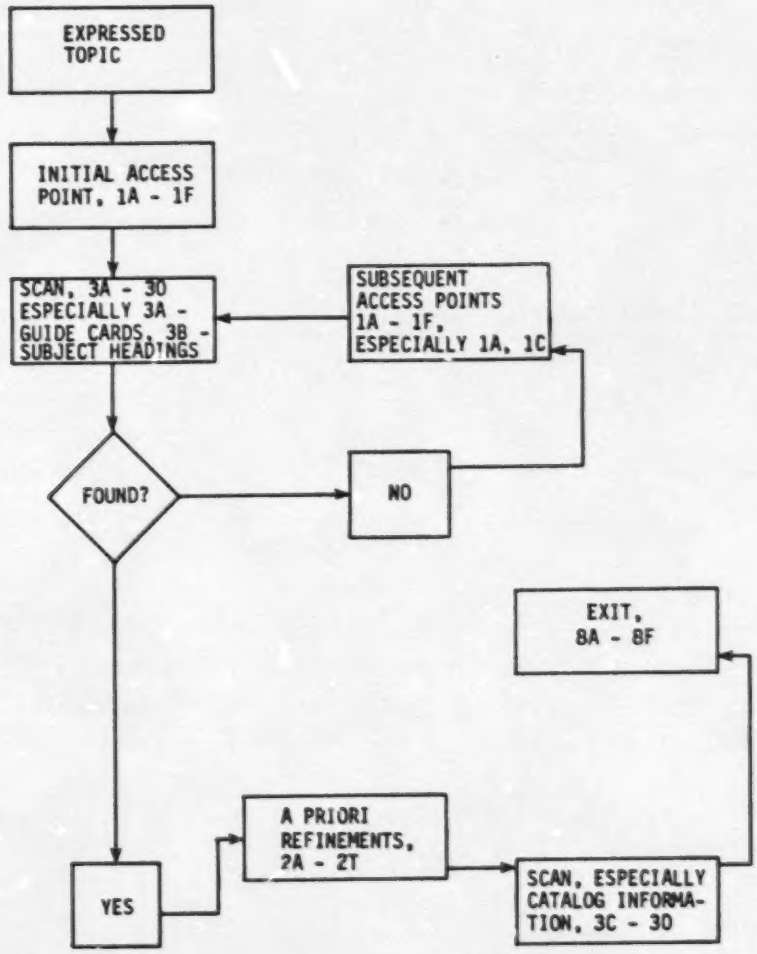

EXPRESSED TOPIC
INITIAL ACCESS POINT, 1A - 1F
SCAN, 3A - 3O ESPECIALLY 3A - GUIDE CARDS, 3B - SUBJECT HEADINGS
SUBSEQUENT ACCESS POINTS 1A - 1F, ESPECIALLY 1A, 1C
FOUND?
NO
EXIT, 8A - 8F
A PRIORI REFINEMENTS, 2A - 2T
YES
SCAN, ESPECIALLY CATALOG INFORMATION, 3C - 3O

Figure 23. Ad Hoc Process Model of a Subject Search for Audi Repair Manual with a priori Refinements

```
PPL> S/AUTOMOBILES, FOREIGN

Call Number   Author / Title     LCCN:73-168562
0             CHILTON BO
629.28        CHILTON'S FOREIGN CAR REPAIR MANUAL(V.02)
C538F         BC:101629251       Monument Branch  Checked Out
Subjects: Automobiles, Foreign - Maintenance and repair.

PPL>F/

Call Number   Author / Title     LCCN:77-095230
RO            MOTOR(NEWY         Reserved, Reserves Office
629.287       MOTOR IMPORTED CAR REPAIR MANUAL(1978)
M919I         BC:103372611       Penrose (Main)
Subjects: Automobiles, Foreign - Maintenance and repair - Collected works.
```

Figure 24. OPAC Search for Automobiles, Foreign in Maggie's Place (the OPAC of Pikes Peak Regional Library District, Colorado)

approximated Boolean "AND" and "OR" operations by scanning the titles assigned the subject heading "Automobiles--Maintenance and repair" for the terms "foreign" or "European." In Chapter III, Section F, such an operation has already been likened to the "E" or Boolean search of Claremont's OPAC and the "stringsearch" of SDC'S ORBIT retrieval software. The OPACs of the University of California (MELVYL), Mankato State University (MSUS/PALS), and Syracuse University (SULIRS) already allow searchers to refine the results of a search argument with the entry and logical combination of the results of a subsequent search argument.

2. Required Features.

A priori refinements are most often associated with subject vocabulary (e.g., keyword, key phrase, or negative keyword or phrase). To effect a priori refinements in the library catalog, searchers must scan items assigned a particular subject heading for the occurrence of the keyword or phrase of interest. OPACs should at least allow subject searchers the capability of refining the results of a previous search argument. Such a capability should not produce irrevocable results so that the searcher has to repeat the initial search argument to undo the damage. The University of California's MELVYL OPAC allows searchers to refine searches by keyword, phrase, or date without losing the results of the initial search argument (Hildreth, 1982b, p. 85). Mankato's OPAC produces intermediate results in Boolean queries, so that searchers can observe the successive decrease in retrieved items as more and more terminology is processed and manipulated by Boolean logic. Furthermore, Boolean operations need not be explicit in OPACs. Rather, the system could prompt the searcher to add refinements to the initial search argument, especially when a large number of postings are retrieved.

A priori refinements occurring in library catalog searches in this study focused primarily on terminology for subject searching (see Chapter III, Section F). Should OPACs allow implicit or explicit Boolean operations, we expect that other types of a priori refinements will become prevalent. These other refinements suggested for OPACs are listed as follows, ranked according to their frequency in subject catalog searches:

1. Publication date
2. Language
3. Dates covered
4. Illustrations
5. Level of specificity
6. Library location

3. Beyond the Library Catalog: A priori Refinements in Subject Searches.

The plight of the library catalog searcher is aptly portrayed in the protocol of Figure 22, wherein failure to match the controlled vocabulary of the subject catalog results in failure to retrieve any items. The searcher's task of finding his terminology in the catalog is complicated by library filing

rules and the interfiling of alphabetically similar terminology. Thus, library catalog searching places the burden of matching and finding terminology on the searcher.

Free text or component word searching of library cataloging records maximizes the OPAC searcher's ability to match the subject terminology of the catalog because matches are allowed on many subject-rich fields in a record. Because of criticism (Cochrane, 1982a, p. 3; Mischo, 1981, p. 22) concerning false drops in search output, alternatives to free text retrieval have been sought.

Mischo offered an approach to maintaining the precoordination of the subject catalog's controlled vocabulary by rotating words in subject headings and titles. "The rotation procedures combined with title keyword/keyphrase access provide a subject retrieval function that, in some measure, simulates a free-text Boolean search capability ... [and] provide greatly increased subject retrieval over the traditional card catalog." (Mischo, 1981, p. 12.)

This approach requires subject searchers to match their terminology with that encased in a single subject heading. It is roughly analogous to a "Field" or "F" text search in Lockheed's DIALOG in controlled vocabulary fields of the MARC unit record. Also comparable is the subject heading browse display of the University of California's MELVYL OPAC (Hildreth, 1982c). Browsing a term such as "Islam" produces a display of subject headings in which "Islam" begins the subject heading (e.g., "Islam--Addresses, essays, lectures") as well as those in which it is encased in the subject heading (e.g., "Christianity and other religions--Islam--Addresses, essays, lectures" or "Communism and Islam") (Cochrane, 1982a, p. 5).

Approaches to the manipulation of subject headings like these of Mischo and the University of California offer a means of maintaining the precoordination of the subject catalog's controlled vocabulary. They also allow access to the innards of subject headings, a capability not present in library catalogs.

D. SUBJECT SEARCHES WITH INTERACTIVE REFINEMENTS

1. Description.

Subject searches containing interactive refinements also contained more than the average number (9.73) of constituent elements and a priori refinements. The searcher typically consulted the card catalog with a partial representation of the expressed topic (1A) or a representation different from the expressed topic (1C). The searcher occupied himself by screening the output (i.e., cataloging records of books) for relevance. In Figure 25, the searcher's protocol discloses that information about finance, statistics, technology, and behavior in connection with multinational corporations is not wanted, whereas general information on multinational corporations prior to 1980, or in connection with Third World countries, is wanted. An ad hoc process model (Figure 26) depicts the subject search for multinational companies and Third World countries. It is very similar to the composite model (Figure 18), but there are no constituent elements for aid and subsequent access points. This model (Figure 26) accommodates the dynamics of the subject search, which waxes and wanes as the searcher scans cards, records

S: I am looking for "Multinational Corporations." So I pull out the one I want to look at and I just go through here till I find Multinational
1A Corporations. As you see, I found it.
DC: As you are flipping through, what is it you are focusing on? What part of the card?
S: I am looking at the top part. The headline, I guess. I am looking for any kind of general information about multinationals because I am doing a thesis on it. So see, there is multinational business and that is what I would look for. I would write that down and then go look for it. What else would you want me to do?
DC: If you would just keep on going through till you find basically what
3C you want. If you see something that interests you and write it down.
S: Do you want to know what I am writing down?
DC: Whatever you would normally write down for your own personal information.
S: Usually I would write down some of the information. Like this is just "Multinational Business," so I know that book is going to cover just mostly business. Then like this one, I might want to use it because it
7F covers the management aspect of the business.
DC: You flipped by a couple of cards that said finance. You are not too
3C interested in that aspect?
2P 4I S: Mostly what I want to cover is just the general aspects of multinational.
2K Like this is "The Businessworld of 1980." I am not covering the 1980s, I am just covering the past. I don't want that. My paper has to focus on the Third World, 'cause I want to do the effects of multinational
2B businesses on the Third World. This is on developing countries so I would like to have this book too. As you can see, most of them are in the Commerce Library. This one is just a multinational company, so just--I need to have information in general on a multinational company--this would be a good book for me too.
DC: You flipped through some of these.
3C S: Well, like "The Multinational Corporation: A Guide in Information
4I Sources." That to me tells me it is mostly a statistic or fact. And I don't want that. This is like Manners and Methods and I don't want
4I anything on that right now. This one is on costs of technology transfer
4N --since my paper doesn't focus on technology transfer; it just focuses on the social--behavioral effects--I don't want that. This multi-
4I national corporation is one I already have. This one is a good one because it has the Problems and Prospects of Multinational Corporations, so that would be another good one.
DC: That is "Multinational Corporation--Social Change."
7F S: That's what I'm looking for. This is just like a general bibliography
3C right now. I am just looking for general stuff right now so I can get
8E started.

Key:

S = Searcher
DC = Data Collector
1A = Access Point, partial representation of expressed topic (i.e., "Multinational Corporations")
3C = Scan, titles
7F = Record, title, author, call number
3C = Scan, titles
2P = A priori Refinement, level of specificity (i.e., "general")
4I = Interactive Refinement, negative keyword (i.e., "Finance")
2K = A priori Refinement, dates covered (i.e., "before 1980")
2B = A priori Refinement, key phrase (i.e., "Third World Countries")
3C = Scan, titles
4I = Interactive Refinement, negative keyword (i.e., "Statistic")
4I = Interactive Refinement, negative keyword (i.e., "Technology")
4N = Interactive Refinement, already have
4I = Interactive Refinement, negative keyword (i.e., "Behavioral")
7F = Record, author, title, call number
3C = Scan, titles
8E = Exit, to another library

Figure 25. Coded Protocol of a Subject Search for Multinational Corporations with Interactive Refinements

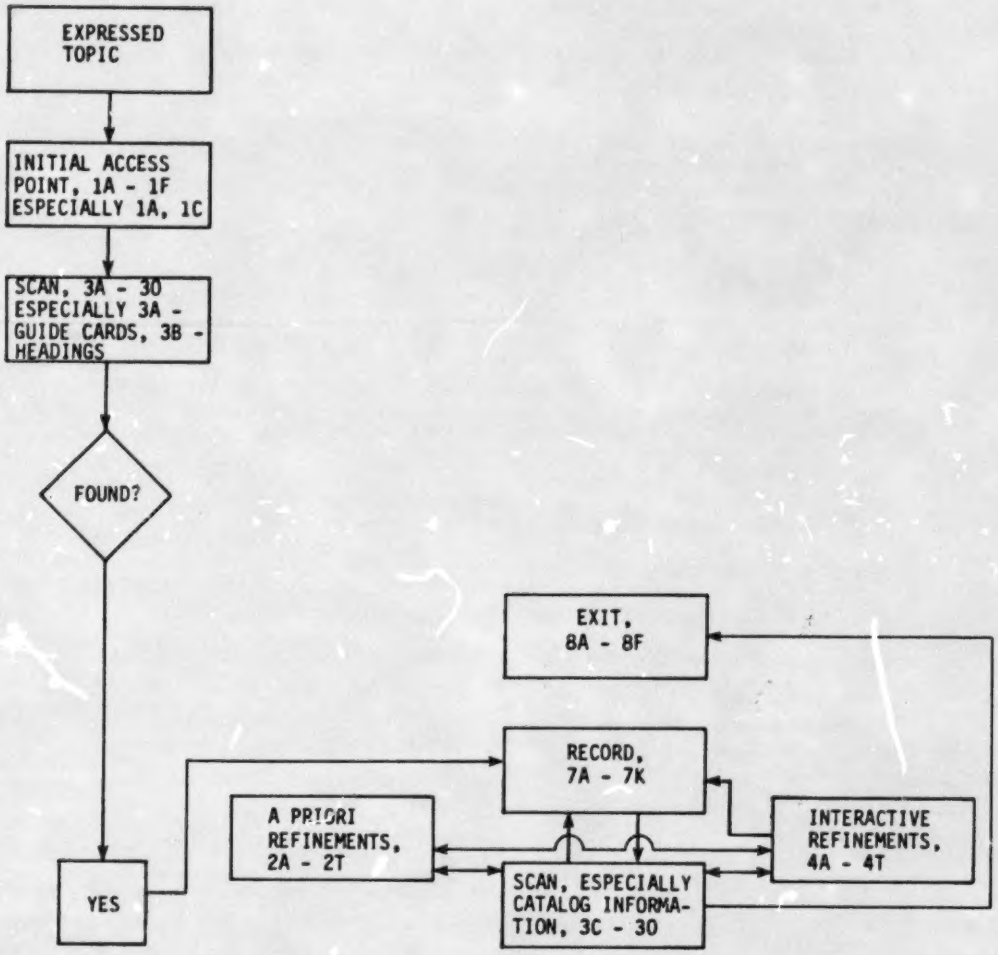

EXPRESSED TOPIC
INITIAL ACCESS POINT, 1A - 1F ESPECIALLY 1A, 1C
SCAN, 3A - 30 ESPECIALLY 3A - GUIDE CARDS, 3B - HEADINGS
FOUND?
YES
EXIT, 8A - 8F
RECORD, 7A - 7K
A PRIORI REFINEMENTS, 2A - 2T
INTERACTIVE REFINEMENTS, 4A - 4T
SCAN, ESPECIALLY CATALOG INFORMATION, 3C - 30

Figure 26. Ad Hoc Process Model of a Subject Search for Multinational Corporations with Interactive Refinements

catalog information, or refines the search on an a priori basis or on the basis of information culled from the catalog. This fluctuation does not occur in a set pattern, especially in the case of interactive refinements, which take place as the search unfolds and catalog information stimulates the searcher to make an on-the-spot refinement.

The most frequent interactive refinements in library catalog searches occurred in connection with subject terminology (see Chapter III, Section F), particularly nonrelevant terminology (4I in Figure 25), which suggests the implementation of the Boolean "NOT" operation in OPACs. Should OPACs allow Boolean operations, we expect that other types of interactive refinements will become prevalent in subject searches. These other refinements suggested for OPACs are listed as follows, ranked according to their frequency in subject catalog searches:

1. Level of specificity
2. Format
3. Illustrations

2. Required Features.

As we have shown, interactive refinements are library catalog searchers' makeshift methods of introducing or deleting subject terminology in the ongoing search strategy. Because of the physical format of the card catalog, refinements are applied to the strategy in a manner similar to the sequential "stringsearch" of SDC's ORBIT retrieval software and the "E" search of Claremont's OPAC. Boolean search capabilities (e.g., those of MELVYL, SULIRS, and MSUS/PALS) could perform the refinement operation in a much more efficient manner than stringsearching. Figures 27 and 28, the two examples of OPAC searches for multinational corporations, show how Boolean operations can produce results with high precision and how differently Boolean capabilities have been implemented in existing online public access catalogs.

The OPAC of Mankato State University (MSUS/PALS) was searched for the topic "Effects of Multinational Corporations on Third World Countries" to show how Boolean operations can pinpoint items with remarkable precision. The record of the search is shown in Figure 27. The "comments" column explains the transaction and shows that much of the dialogue (Boolean operators, save sets commands, display commands, etc.) was initiated by the searcher.

In contrast, the Boolean "AND" operation in Syracuse University's OPAC (SULIRS) is activated implicitly when searchers enter a string of subject terms on the same line. Figure 28 shows the subject search of SULIRS. Although its display is "cleaner" (less verbose) than the Mankato OPAC, the inclusion of subject headings in the short format records of Figure 28 could have helped the searcher to find out more information about the books' subject contents or to find additional search terminology, particularly in this search where none of the terminology of the expressed topic matches a subject heading verbatim.

Refinements in library searches are often connected with accessing, scanning, and recording activities. Thus, Boolean operations in OPACs may occur at

ONLINE OPAC SEARCH IN MSUS/PALS	COMMENTS
>TE MULTINATIONAL CORPORATIONS ANDS WERE ASSUMED FOR MISSING BOOLEAN OPERATORS 141 RECORD MATCHES AFTER TERM MULTINATIONAL 47 RECORDS MATCHED THE SEARCH TYPE DI 1-20 TO DISPLAY FIRST 20 RECORDS >SAVE WORK AREA SAVED IN SET 1 >TE DEVELOPING COUNTRIES ANDS WERE ASSUMED FOR MISSING BOOLEAN OPERATORS 10 RECORDS MATCHED THE SEARCH TYPE DI 1-10 TO DISPLAY FIRST 10 RECORDS >DI 1 LONG	"Multinational corporations" is retrieved from the title field of the MARC record. The 47 items matching the search are saved for future processing. To find the "correct" subject heading, items whose titles or subject terms contained "developing countries" were retrieved. The first item was displayed to find the "correct" subject heading for "developing countries."
Screen 001 of 001 Record 0001 of 0010 MSU LOCTN: RA427 .047 1975 AUTHR: Omran, Abdel R. TITLE: Community medicine in developing countries. Abdel R. Omran, editor. Participating Advisory Board: John C. Cassel [and others] PUBLR: New York, Springer Pub. Co. [1975, c 1974] DESCR: xxiii, 507 p. illus. 24 cm. NOTES: Includes bibliographies. SUBJT: Public health. SUBJT: Developing countries. SUBJT: Community health services. SUBJT: Community health services. SUBJT: Underdeveloped areas--Public health. OCLC# 00969549	
>SAVE WORK AREA SAVED IN SET 2 >TE UNDERDEVELOPED AREAS ANDS WERE ASSUMED FOR MISSING BOOLEAN OPERATORS 589 RECORD MATCHES AFTER TERM UNDERDEVELOPED 571 RECORDS MATCHED THE SEARCH >SAVE WORK AREA SAVED IN SET 3	"Underdeveloped areas," a subject heading, will now be incorporated into the strategy.
>BO 2 OR 3 10 RECORD MATCHES AFTER SET 2 574 RECORDS MATCHED THE SEARCH >SAVE WORK AREA SAVED IN SET 4	Since "underdeveloped areas" and "developing countries" are somewhat synonymous, saved sets for both were combined in a Boolean "or" operation.
>BO 1 AND 4 47 RECORD MATCHES AFTER SET 1 3 RECORDS MATCHED THE SEARCH	The combination of sets for "multinational corporations" (set 1) and "underdeveloped areas" (set 4) results in 3 titles.

Figure 27. OPAC Search for Multinational Corporations in Mankato State's MSUS/PALS Featuring Explicit Boolean Operations

ONLINE OPAC SEARCH IN MSUS/PALS

```
>DI LONG 1-3
Screen 001 of 001     Record 0001 of 0003 MSU
LOCTN: REFERENCE Z674.5 .C6x no.745
AUTHR: Steele, R      D
TITLE: Multinational corporations, technology transfer, and the developing
       countries; an introductory bibliography.
PUBLR: Monticello, Ill., 1975.
DESCR: 39 p. 28 cm.
SERIE: Council of Planning Librarians. Exchange bibliography; no. 745
NOTES: Cover title.
SUBJT: Underdeveloped areas--Bibliography.
SUBJT: International business enterprises--Bibliography.
SUBJT: Corporations, International--Bibliography.
OCLC#  01185852
----Type NR to proceed thru display list
>NR
Screen 001 of 001     Record 0002 of 0003 MSU
LOCTN: HD2755.5 .S65
AUTHR: Solomon, Lewis D.
TITLE: Multinational corporations and the emerging world order / Lewis D.
       Solomon.
PUBLR: Port Washington, N.Y. : Kennikat Press, 1978.
DESCR: 261 p. ; 22 cm.
NOTES: Includes bibliographical references and index.
SUBJT: Food and Agriculture Organizaion of the United Nations. Industry
       Cooperative Programme.
SUBJT: International economic relations.
SUBJT: Underdeveloped areas--Investments, Foreign.
SUBJT: International business enterprises.
OCLC#  03447363
----Type NR to proceed thru display list
>NR
Screen 001 of 001     Record 0003 of 0003 MSU
LOCTN: HD2755.5 .W67x
TITLE: The world as a company town : multinational corporations and social
       change / editors, Ahamed Idris-Soven, Elizabeth Idris-Soven, Mary
       K. Vaughan.
PUBLR: The Hague ; Paris : Mouton ; Chicago : distributed in the USA and
       Canada by Aldine, c1978.
DESCR: xi, 455 p. ; 24 cm.
SERIE: World anthropology
NOTES: "...the result of a panel on multinational corporations held at the
       IXth International Congress of Anthropological and Ethnological
       Sciences."
NOTES: Includes bibliographies and indexes.
SUBJT: Underdeveloped areas--International business enterprises--Congresses.
SUBJT: International business enterprises--Social aspects--Congresses.
SUBJT: International business enterprises--Congresses.
AAUTH: Vaughan, Mary K., ed.
AAUTH: Idris-Soven, Elizabeth, ed.
AAUTH: Idris-Soven, Ahamed, ed.
AAUTH: International Congress of Anthropological and Ethnological Sciences,
       9th, Chicago, 1973.
OCLC#  04828433
```

COMMENTS

Full bibliographic records of the 3 retrieved items are displayed. Subject headings such as "International business enterprises" and "Corporations, International" could be incorporated into the strategy.

Interestingly, two of the three items have been classed in HD2755.5. The searcher could browse the bookshelves online if there were a call number access point capability in MSUS/PALS.

Figure 27.--Continued

ONLINE OPAC SEARCH IN SULIRS

```
ENTER SEARCH: WD;MULTINATIONAL CORPORATIONS UNDERDEVELOPED AREAS
SEARCHING "MULTINATIONAL"
SEARCHING "CORPORATIONS"
SEARCHING "UNDERDEVELOPED"
SEARCHING "AREAS"
   9 ITEMS.  PRESS RETURN FOR DISPLAY, OR S AND RETURN FOR NEW SEARCH:

   1. FEDER, ERNEST
      THE NEW PENETRATION OF THE AGRICULTURES OF THE UNDERDEVELOPED COUNTRIES BY
      THE INDUSTRIAL NATIONS AND THEIR MULTINATIONAL CONCERNS
      GLASGOW, UNIVERSITY OF GLASGOW, 1975
      HD1417.F43 ; 4TH FLOOR-BIRD ; MISSING, BEING TRACED

   2. SOLOMON, LEWIS D
      MULTINATIONAL CORPORATIONS AND THE EMERGING WORLD ORDER
      PORT WASHINGTON, N.Y. KENNIKAT PRESS, 1978
      HD2755.5.S65 ; 4TH FLOOR-BIRD ; IN CIRCULATION - DUE 05/22/82

   3. GUPTA, N. S
      MULTINATIONAL CORPORATIONS : A STUDY OF SOCIO-ECONOMIC IMPLICATIONS ON THE
      COUNTRIES OF THIRD WORLD : 1ST ED, MEERUT, PRAGATI PRAKASHAN, 1978
      HD69.I7G87 ; 4TH FLOOR-BIRD ; MISSING, BEING TRACED

   9 ITEMS.  PRESS RETURN FOR DISPLAY, OR S AND RETURN FOR NEW SEARCH:
```

```
   4. IDRIS-SOVEN, AHAMAD, 1947- *
      THE WORLD AS A COMPANY TOWN : MULTINATIONAL CORPORATIONS AND SOCIAL CHANGE
      THE HAGUE, MOUTON, C1978
      HD2755.5.W67 ; 4TH FLOOR-BIRD

   5. SEN, SAMARENDRA CHANDRA, 1922-
      MULTINATIONAL CORPORATIONS IN THE DEVELOPING COUNTRIES
      CALCUTTA, EASTERN LAW HOUSE, 1978
      HD2755.5.S45 ; 4TH FLOOR-BIRD

   6. LA PALOMBARA, JOSEPH G *
      MULTINATIONAL CORPORATIONS AND DEVELOPING COUNTRIES : A RESEARCH REPORT FROM
      THE CONFERENCE BOARD ; NEW YORK, CONFERENCE BOARD, C1979
      HD69.I7L35 ; 4TH FLOOR-BIRD ; IN CIRCULATION - DUE 09/15/82

   •
   •
   •
```

COMMENTS

The "WD" search retrieves items with the specified terms in the titles or subject headings of SULIRS records.

Three of the six items displayed were classed in HD2755.5. The searcher could now use SULIRS' "CN" or call number search to browse the shelf list online.

Figure 28. OPAC Search for Multinational Corporations in Syracuse's SULIRS Featuring Implicit Boolean Operations

times when such activities are going on. Hildreth's (1982b, pp. 65-66) notion of "suggestive prompts to inform the user of possible action(s) which may or may not be taken" may help OPAC searchers to incorporate Boolean operations into their strategy and cut down nonrelevant output. For example, access points which result in 200 or more retrieved items in OPACs could be accompanied by a "suggestive prompt" which asks the searcher whether he would like to refine the retrieved items by entering additional search terminology. Likewise, when the searcher scans through displays of 100 or more items, a suggestive prompt could be introduced into the dialogue which asks the searcher what nonrelevant terminology continually appears in the output, or what terminology has (or has not) appeared in the online display that is particularly interesting. Searchers' responses to suggestive prompts would activate the appropriate Boolean operation, which would be transparent to searchers. Our studies have shown that the initial access point in library catalog searches rarely exceeds two terms and typically involves a partial representation of searchers' expressed topics (see Chapter III, Section D). Suggestive prompts may be a means of eliciting the "other part" of the expressed topic from searchers and effecting more precise output.

3. Beyond the Library Catalog: Interactive Refinements in Subject Searches.

The searcher who looked for "Multinational Corporations" in a subject search of the library catalog (Figure 25) was fortunate that he was searching a dictionary catalog, which interfiles authors, titles, and subject headings. Thus, his chances of matching his terminology with that of the catalog increased because the dictionary catalog permits access to the first words of titles. Luckily, he found titles beginning with "Multinational;" the "correct" subject heading for representing "Multinational corporations" is "International business enterprises." The subject search capability of Mankato State's MSUS/PALS and Syracuse's SULIRS allows retrieval of all terms in titles and subject headings. This is not possible in the traditional subject catalog. Access to the innards of titles and subject headings through the University of California's subject heading browse command or Mischo's rotated title words and subject headings maintains the precoordination of titles and subject headings and is an alternative to free text search and retrieval of subject-rich MARC fields through Boolean operations.

Another approach to maintaining the precoordination of subject headings would be to obtain feedback from the OPAC searcher about relevant output. CITE, one prototype OPAC of the National Library of Medicine (Doszkocs, 1981), incorporates such feedback into the online dialogue and ranks output on a best-match-of-searched-terms basis. As CITE and OPACs with comparable feedback mechanisms develop, we can hope for feedback which ranks output on a best-match-of-searched-facets basis (i.e., the first items listed contain terms representing all three facets, then two of three facets, and so on) and automatically suggests to searchers the inclusion of controlled vocabulary when there are very few or no postings for a term.

E. SUBJECT SEARCHES, MULTIPLE ACCESS POINTS, AND ASSISTANCE

1. Description.

Subject searches with multiple access points are marked by the searchers' inability to find the exact subject terminology to represent their expressed topic. The protocol for a subject search on FCC (Figure 29) contains three access points--U.S. Government, Media, and Mass media. The subject searching process for FCC is depicted in Figure 30. Missing from this model, in comparison with the composite model (Figure 18), are records, refinements, and librarian assistance. The search apparently failed, since no items were recorded and no mention was made of reading FCC in the catalog. If only the searcher had consulted a librarian who could have suggested just the right subject heading for FCC. The searcher encountered the only means of guidance for terminology in the subject catalog itself (i.e., the cross reference).

Librarian assistance was sought in a subject search for "Life after Death" (Figure 31). Just like the searcher's plight when searching for FCC, the searcher of "Life after Death" could not find the exact subject heading to express his topic. The librarian's suggestion of "Future life" resulted in a promising title, but the searcher seemed to be looking for titles which contained his terminology verbatim (i.e., "Life after Death") so that the expressed topic was like an a priori refinement in this search.

The ad hoc model of the subject searching process for "Life after Death" is almost a full-blown representation of the composite model save the absence of a priori refinements and paths from a number of activities to exit or subsequent access points (Figure 32). Dotted lines in the model denote paths which are frequent in searches of this type but which were not taken in this example.

2. Required Features.

The key to success at the subject catalog and the online public access catalog is the matching of the searcher's expressed topic with the catalog's controlled vocabulary. The two searchers who expressed failure at the catalog (Figures 29 and 31) both selected terminology which matched neither their expressed topic nor the catalog's controlled vocabulary. Yet both encountered two sources of assistance provided by the library and the library catalog (i.e., cross references and reference assistance). The Library of Congress Subject Headings (LCSH), another source of assistance, was not consulted by the searchers.

Hildreth's (1982b) study of ten operational OPACs disclosed that no OPACs presently provide cross references for subject headings, although two systems (the University of California's MELVYL and The Ohio State University's LCS) plan to implement this feature in the near future. MELVYL provides a cross-reference capability for author names, but fails to report that it has also retrieved items written by the same author requested by the searcher but published under a different name. Hildreth (1982b, p. 119) recommends that OPACs either obtain the searcher's permission or inform the searcher when retrieving cross references, particularly "see references," so that the

S: I'm looking for information on the FCC.
DC: And you are looking under U.S.?
S: Yes
1C DC: "U.S. Government"?
S: Yeah.
DC: What part of the card are you reading? Are you just looking at the top line?
S: Yes, basically.
3B DC: What else are you reading?
3C S: Well, sometimes they give information about what the book is about.
DC: Would you just read out loud what you are reading?
S: Just looking at top lines really.
DC: The stuff in red?
S: Yeah. I doubt if it's in here. I don't think they got it.
DC: What are you going to look under now?
1C S: "Media."
3G DC: Looked under "Media" and it said to see what--"Mass Media"? Trying
1C to get back to the beginning of this? What part are you reading now?
S: The description of books.
DC: Just reading the titles or just the first few lines?
3C S: Just the first few lines ... "Privacy in the Press." Well that's
8H about it.

Key:

S = Searcher
DC = Data Collector
1C = Access Point, different from expressed topic (i.e., "U.S. Government")
3B = Scan, subject headings
3C = Scan, titles
1C = Access Point, different from expressed topic (i.e., "Media")
3G = Scan, cross-reference (i.e., "Media, see Mass Media")
1C = Access Point, different from topic (i.e., "Mass Media")
3C = Scan, titles
8H = Exit, unknown

Figure 29. Coded Protocol of a Subject Search for FCC with Multiple Access Points and No Records (i.e., Failure)

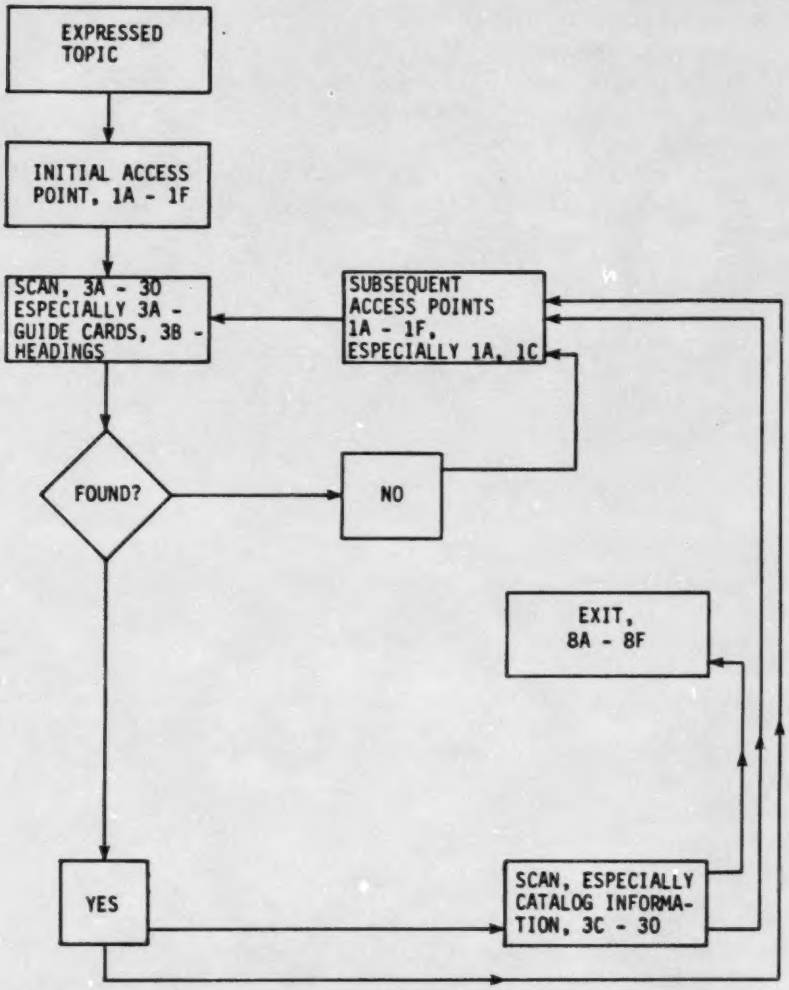

EXPRESSED TOPIC
INITIAL ACCESS POINT, 1A - 1F
SCAN, 3A - 3O ESPECIALLY 3A - GUIDE CARDS, 3B - HEADINGS
SUBSEQUENT ACCESS POINTS 1A - 1F, ESPECIALLY 1A, 1C
FOUND?
NO
EXIT, 8A - 8F
YES
SCAN, ESPECIALLY CATALOG INFORMATION, 3C - 3O

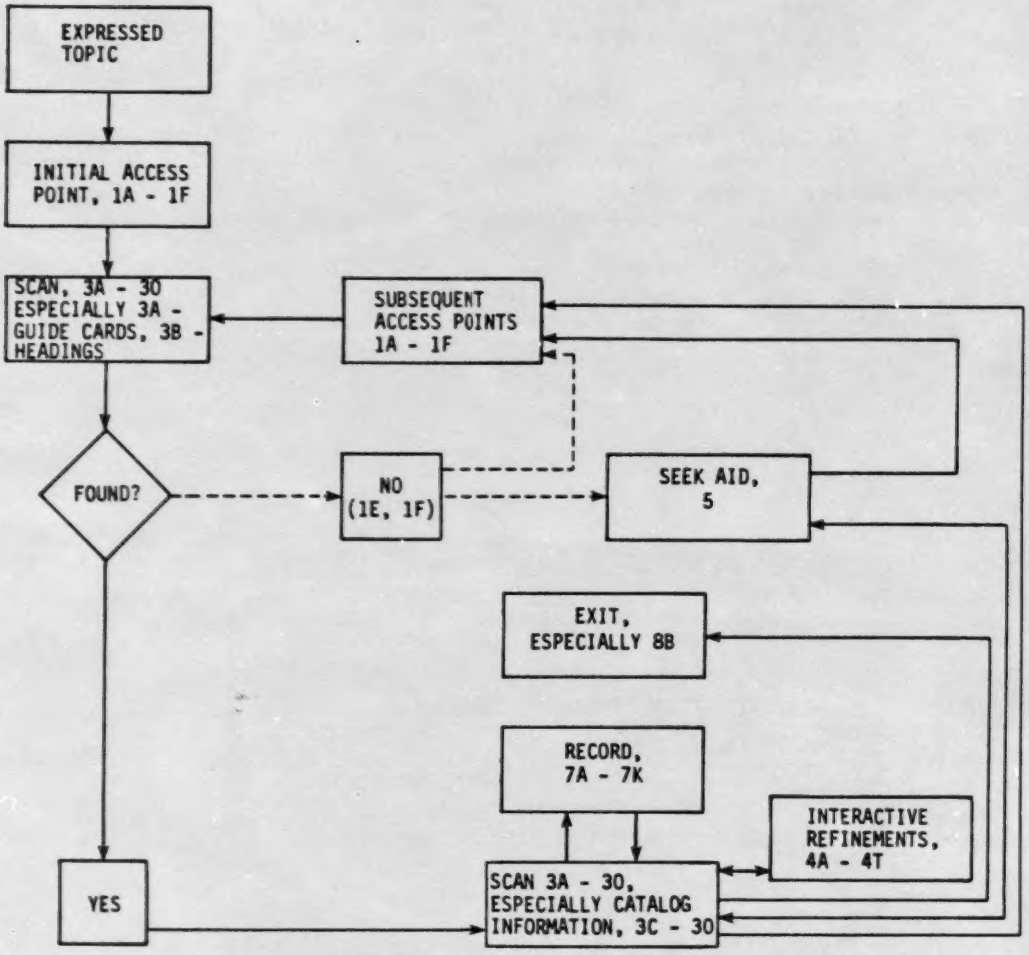
Figure 30. Ad Hoc Process Model of a Subject Search for FCC with Multiple Access Points and No Records (i.e., Failure)

1A S: Life.
DC: OK. You're looking for life after death, right?
S: It might be too hard. I just want to see what I can find on it .... I don't think there is anything on it.
3B DC: "Life Science," "Life Spirit," "Life Saving."
5 Library Aide: You might try "Mortality."
DC: She suggested looking under Death when she started out? Now which part of the card are you looking at, the first lines?
S: No, the next line down.
3C DC: OK. So it's the title. Why did you not like that one?
5,1C Library Aide: Try "Future Life."
DC: Now are you just looking at the first lines?
3J S: Now I'm looking at everything--but it just says "death and dying."
7H I might write it down.
DC: Those are just about death and stuff.
S: Maybe that one ...
3C DC: "The Hour of Death"? It's fiction.
4I,3C S: No. Darn! There's another one here too .... She (library aide) suggested "Immortality." That's the same one we had before. That's the only one
1C,3C,8A there is.

Key:

S = Searcher
DC = Data Collector
1A = Access Point, partial representation of expressed topic (i.e., "Life")
3B = Scan, subject headings
5 = Librarian assistance
3C = Scan, titles
1C = Access Point, different from expressed topic (i.e., "Future Life")
3J = Scan, full bibliographic information
7H = Record, unknown
3C = Scan, titles
4I = Interactive refinement, negative keyword (i.e., "Fiction")
3C = Scan, titles
1C = Access Point, different from topic (i.e., "Immortality")
8A = Exit, to stacks

Figure 31. Coded Protocol of a Subject Search (with Librarian Assistance) for Life after Death


EXPRESSED TOPIC
INITIAL ACCESS POINT, 1A - 1F
SCAN, 3A - 30 ESPECIALLY 3A - GUIDE CARDS, 3B - HEADINGS
SUBSEQUENT ACCESS POINTS 1A - 1F
FOUND?
NO (1E, 1F)
SEEK AID, 5
EXIT, ESPECIALLY 8B
RECORD, 7A - 7K
INTERACTIVE REFINEMENTS, 4A - 4T
YES
SCAN 3A - 30, ESPECIALLY CATALOG INFORMATION, 3C - 30

(Note: Dotted line indicates process which frequently occurs but did not occur in the search modeled.)

Figure 32. Ad Hoc Process Model of a Subject Search (with Librarian Assistance) for Life after Death

searcher is not confused by the system's apparent independent action. Underscoring the need for cross references in OPACs, he also advises that "Online catalogs that do not provide 'see' and 'see also' references for term selection deprive users of aids that they may already have grown accustomed to (and benefitted from) when searching the card catalog." (Hildreth, 1982b, p. 121.)

Not one of the 189 searchers consulted the LCSH during their subject catalog searches and only 15 sought the librarian's assistance, although their reticence to consult the LCSH or the librarian might have been due to the method of data collection in this study. Since the data collector followed searchers and prompted them to give voice to their thoughts, the inclusion of another person or written source in a usually solitary activity might have been awkward. Searchers' problems with selection of controlled vocabulary will undoubtedly continue after OPACs are introduced, so that the LCSH and librarians should be nearby or accessible to OPAC searchers.

3. Beyond the Library Catalog: Multiple Access Points and Assistance in Subject Searches.

The capability of placing the OPAC in locations other than the library or its branches (e.g., dormitories and university or city offices) and accessing it from remote locations via cable television or dial-in access underscores the need for assisting the searcher while he is using the online public access catalog. Online aids for finding, browsing, and selecting controlled vocabulary are imperative. The related term display in Figure 33 was compiled from the "see also" references of the printed LCSH; the class numbers which accompany subject headings in the printed LCSH are included. The number of postings were taken from searches of the subject headings in the OPAC of The Ohio State University. The display should not automatically be activated when the searcher types in the term "Islam." Rather, the system should immediately report the number of "hits" (e.g., 68) and include a suggestive prompt to inform the searcher that a display of terms related to "Islam" can be produced. Note that the display includes class numbers so that searchers can go directly to the bookshelves.

Many subject searchers withhold the vocabulary of their expressed topic and access the card catalog with either a partial representation of or terms different from their expressed topic. Rarely do searchers' access points exceed two words. Online vocabulary displays might be a means of stimulating searchers to respecify their access point through the vocabulary of online displays because of the variety and complexity of the controlled vocabularies in online displays. For example, the University of California's MELVYL produces a keyword-in-context controlled vocabulary display. Searchers are able to retrieve all subject headings in which their search term is encased. Such an operation is virtually impossible using the traditional library catalog and printed LCSH; even to attempt it requires time-consuming effort and "detective" work. MELVYL's keyword-in-context display for "Islam" is depicted in Figure 34. Unfortunately, there are no helpful prompts concluding the display to guide the searcher to the next step. The mere length of the display could frighten searchers. Cochrane (1982a) suggests that controlled vocabulary displays such as the keyword-in-context example be categorized by subdivisions (e.g., geographical, chronological, or format). Figure 35 is an example of how an online display of Islam subject headings with geographical

User: RBL/ISLAM

RBL	BOOKS	SUGGESTED CLASS NO.	RELATED SUBJECTS
1	68	BP1-BP223	Islam
2	5		Caliphate
3	4	BP175.D4	Dervishes
4	1	BP166.94	Imamate
5	2		Jerusalem in Islam
6	1		al-Jum'ah
7	5	BP100-BP130	Koran
8	1	BP188.16.L68	Love (Islam)
9	1	BP166.93	Mahdism
10	3	BP188.16.M6	Modesty (Islam)
11	1	BP189.33	Muslim Saints
12	4	DS38	Muslims
13	1		Palestine in Islam
14	3	BP166.4	Prophets, Pre-Islamic
15	1	BP190.5.S4	Sex and Islam
16	1	HV6065	Temperance and Islam
17	2	BP190.5.W35	War and Islam

PAGE 1 OF 1 FOR TITLES ON A SPECIFIC SUBJECT ENTER TBL/ AND LINE NUMBER

Figure 33. Related Term Display with Class Areas Constructed from the Printed Library of Congress Subject Headings

Browse request: BROWSE SU ISLAM
Browse result: 254 subject headings found in the subject index

1. Authority (Islam)
2. Bibliography -- Bibliography -- Islam
3. Birth control -- Religious aspects -- Islam
4. Birth control -- Religious aspects -- Islam -- Congresses
5. Blasphemy (Islam)
6. Christianity and other relgions -- Islam
7. Christianity and other religions -- Islam
8. Christianity and other religions -- Islam -- Addresses, essays, lectures

•
•
•

33. God (Islam) -- Proof
34. Good and evil (Islam)
35. Hinduism -- Relations -- Islam
36. Intermediate state (Islam)
37. Islam
38. Islam (The Arabic word)
39. Islam -- Addresses, essays, lectures
40. Islam -- Africa
41. Islam -- Africa -- Addresses, essays, lectures
42. Islam -- Africa -- Bibliography
43. Islam -- Africa -- Congresses
44. Islam -- Africa, East

•
•
•

230. Science -- Islam -- Bibliography
231. Sex and Islam
232. Sex and Islam -- Congresses
233. Sikhism -- Relations -- Islam
234. Sin (Islam)
235. Slavery and Islam -- Africa, East
236. Socialism and Islam

•
•
•

253. Tabrizi, Sigat al-Islam Shahid, 1860-1911
254. Vereeniging Sarekat Islam

Figure 34. Keyword-in-Context Display of Subject Heading (Islam) from the University of California's MELVYL

```
>Searcher:  BROWSE SU ISLAM
254 subjects found.  Subjects can be listed with:

        1.  type of materials, e.g., Bibliography.
        2.  geographic areas, e.g., East Africa.
        3.  time periods, e.g., Middle Ages.
        4.  other subjects, e.g., Sex.
        5.  all of the above


Which line above is closest to your interest?

>2
```

Screen 1

```
 1.  Communism and Islam--Sudan
 2.  Fasts and feasts--Islam--Indonesia--Jakarta
 3.  Fasts and feasts--Islam--Indonesia--Medan
 4.  Islam--Africa
 5.  Islam--Africa--Addresses, essays, lectures
 6.  Islam--Africa--Bibliography
 7.  Islam--Africa--Congresses
 8.  Islam--Africa, East
 9.  Islam--Africa--History
10.  Islam--Africa, North
 .
 .
 .
22.  Islam--Egypt
23.  Islam--Egypt--History
24.  Islam--Ethiopia
25.  Islam--Gujarat
 .
 .
 .
39.  Islam--Upper Volta
40.  Islam--Zaire
 .
 .
 .

Which lines above are closest to your interest?
```

Screen 2

Figure 35. Keyword-in-Context Display of Subject Heading (Islam) with Geographical Subdivisions

subdivisions would be activated in an OPAC search. The addition of postings and class number areas in both keyword-in-context displays (Figures 34 and 35) would assist bookshelf browsers.

Postcoordination of subject and title words through Boolean operations has already been discussed as a means of refining the output of subject searches and increasing precision. It could also be connected to a feature which tracks the subject headings and class number areas in the output and suggests the same to the searcher based upon subject headings and class number areas common to the greatest number of items in the output (Atherton and Markey, 1979). For example, an OPAC search in Syracuse's SULIRS for "FCC" retrieved eight items. Class number areas and subject headings assigned to the eight items and the number of times they occur are:

No. of Items	Subject Headings and Class Number Areas
3	United States. Federal Communications Commission
3	TK6554.5.S9
2	Broadcasting Policy--United States
2	HE8689.8.C56
1	Radio--Examinations, questions, etc.
1	Radio Operators--United States
1	AS36.R282R1896MF
1	KF2805.A2W44
1	KF2844.L4

Two subject headings and two class number areas are common to five of the eight retrieved items. A suggestive prompt following the display of the eight titles would inform the searcher of subject headings and class number areas in which he might also be interested (Figure 36).

PaperChase, the retrieval system for journal citations at Beth Israel Hospital, already features a link between title terms and Medical Subject Headings (MeSH) and prompts the searcher when a controlled vocabulary term is available to express the concept (Horowitz and Bleich, 1981; Cochrane, 1982b). Mischo's idea of rotating title terms in MARC records and producing subject heading displays in response to title word searches is comparable but would not allow searchers to obtain the results of title keyword searches alone.

Cochrane (1982a, p. 5) also envisions other features to assist searchers in the selection of vocabulary for subject searching in OPACs:

- Automatic interaction between searcher and computer system so that the best list of search terms can be found.
- Automatic display of some retrieved items for relevance, feedback, and iteration of the above feature using different selection rules.
- Automatic display of logical outlines (or classified lists) to select search terms in context: some broader, some coordinate, some narrower. This may help increase the output or sharpen the precision of the search.

```
                •
                •
                •
                8.
Call Number:    HE8689.8.C56
Author:         COLE, BARRY G
                OETTINGER, MAL, 1932-
Title:          RELUCTANT REGULATORS : THE FCC AND THE BROADCAST AUDIENCE
Publisher:      READING, MASS, ADDISON-WESLEY PUB. CO/C1978
Language        ENGLISH
Availability:   DISCHARGED - ON 4/15/82
Subjects:       1. UNITED STATES. FEDERAL COMMUNICATIONS COMMISSION
                2. BROADCASTING POLICY--UNITED STATES

Try these subjects and class numbers for more books:

   1.   United States. Federal Communications Commission
   2.   TK6554.5.S9
   3.   Broadcasting Policy--United States
   4.   HE8689.8.C56

Which lines above are closest to your interest?
```

Figure 36. Online Vocabulary Aid for Suggesting to OPAC Searchers Subject Headings and Class Number Areas Common to Displayed Items

```
SEARCHER:  Browse ISLAM/CT Summary

           297 ISLAM and religions derived from it

           Summary

   297.1   Sources, relationship, attitudes of Islam
      .2   Islamic doctrinal theology (Aqaid and Kalam)
      .3   Islamic public worship and other practices
      .4   Islamic religious experience, life, practice
      .5   Islamic moral theology
      .6   Islamic leaders and organization
      .7   Islamic activities
      .8   Islamic sects and other religions derived from Islam

Which line above is closest to your search topic???_____
```

(Source: Cochrane, 1982c, "Figure 1. Screen Display with Classification Summary")

Figure 37. The Dewey Decimal Classification Scheme in the Online Public Access Catalog

- Automatic display of limit options (language, date of publication, etc.). This could help reduce the output of the search.

Cochrane's call for logical outlines or classified lists is presently available only through OPAC searches using truncated call numbers. This feature simulates searching the library's shelflist. CITE, one of the prototype NLM OPACs, incorporates class numbers into its feedback routine. However, the way in which class number searching is done in an OPAC requires searchers to know the meaning of the class number. In the late 1960s, Freeman and Atherton (1968) experimented with information retrieval through automatic manipulation of the Universal Decimal Classification (UDC) schedules in the AUDACIOUS Project. In view of their findings, Cochrane (1982c, 1982d) now places research in classification into the context of the online public access catalog and has initiated further research to explore the subject retrieval potential from the text (i.e., classification captions of the Dewey Decimal and Library of Congress Classification (DDC and LCC, respectively) schedules). Screen displays of the DDC show how the classification's hierarchical structure could lead searchers to topics which are coordinate or broader or narrower than the vocabulary used to represent their access point (Figure 37). Such hierarchical displays of controlled vocabulary or classification captions are necessary to draw out the expressed topics from searchers, since, as described in detail, they are prone to consult a partial, different, or shortened representation of their expressed topic.

Online vocabulary displays can assist OPAC searchers when selecting access points. Such displays were never possible in traditional library catalogs. Technology makes it possible for an OPAC to manipulate the Library of Congress Subject Headings in many ways to stimulate OPAC searchers. An OPAC can provide:

1. Cross references to preferred subject headings activated by the searcher following a suggestive prompt.

2. Related subject heading displays like the printed arrays in LCSH.

3. Keyword-in-context displays of subject headings, such as the University of California's MELVYL already provides.

4. Keyword-in-context displays of subject headings by type of subdivision (e.g., geographical or chronological).

5. Automatic detection of subject headings and class number areas common to the greatest number of displayed or retrieved items.

6. Logical outlines or lists from classification schedules to put topics in context.

F. THE USER-ORIENTED INTERFACE IN ONLINE PUBLIC ACCESS CATALOGS

As data collectors, we observed high school students and public, college, and university library users in the process of subject searching the library catalog. Although searchers approached the library catalog and performed

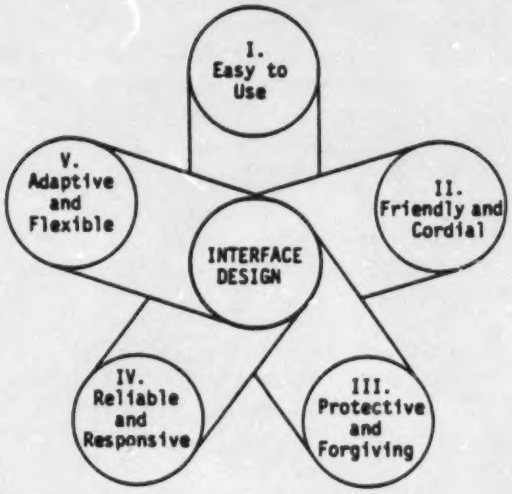
their search with varying degrees of proficiency, none was stumped or bewildered by the intellectual exercise of library catalog searching or overwhelmed by the physical skills required for button, card, guide card, or drawer manipulation. Physical and intellectual activities necessary for traditional library catalog searching are exercised nearly every day when we scan and select files from a file cabinet, browse a recipe file, or look for either a phone number in the telephone directory or the correct spelling or meaning of a word in the dictionary. The same skills that we use to search the catalog at our neighborhood public library are used when we search the catalog of a major research library. Thus, library catalog searching displays a certain constancy both in terms of its physical nature and of our possession of adequate physical skills to perform it.

In contrast, there are aspects of the online public access catalog which have never been encountered by searchers of the library catalog and will be wholly new to those searching OPACs. Many OPAC searchers do not realize that the catalog, instead of residing in the computer terminal, is electronically transmitted to and displayed on it in response to searchers' messages. Presently, those messages are not only activated by the searchers' touch on a typewriter-like keyboard, function keys, or heat-sensitive screens, but they also must be tailor-made to accomplish the specific function (e.g., display a specific item, display a range of items, search for a topic, or browse subject headings) and to communicate to the particular OPAC (e.g., display 1, find su islam). The constancy of the library catalog cannot be carried over to the online public access catalog, primarily because the physical skills inherent in library catalog searching must be translated or "logicalized" into messages that are understood by searcher and computer alike.

For example, when searching and locating subject headings in a card catalog, searchers "naturally" flip through the first four or five cards to determine if the items assigned to the heading are relevant to their topic. To accomplish this same task in an OPAC, the user must specify what he wants done (e.g., search, display, print, type), what format is to be shown (e.g., full record, titles only, brief record), and how many he wants done (e.g., all, one at a time, first five only). At this time, function names, types of display formats, and the number of records that can be displayed at one time vary from OPAC to OPAC (Hildreth, 1982b; Cochrane, 1982a).

The constancy of the library catalog--the very quality that contributes to its ease of use, familiarity, and unimposing nature--is virtually impossible to build into the OPAC because the medium of the OPAC precludes the physical activities inherent in library catalog searching. Thus, we cannot rely on our models of traditional catalog searching when incorporating the constancy of the library catalog into the online public access catalog.

> As interactive computer systems, online catalogs add an entirely new dimension to catalog use not present or encountered in noninteractive catalog systems. In terms of process, this new dimension is referred to as human-computer interaction or man-machine communication. In structural terms it is known as the user-system interface or man-machine interface component of the computer system or environment. The user-system interface, or simply, user interface, is the locus of that interaction undertaken in the use of interactive computer systems, such as online catalogs and other information retrieval systems. (Hildreth, 1982b, p. 33; emphasis added.)



I. Easy to Use

- Easy to access and begin use
- Easy to understand basic functions/capabilities
- Easy to learn command use
- Easy to enter commands/data
- Easy to grow, advance, in use
- Easy to remember how to use over time
- Easy to refresh/relearn how to use

II. Friendly and Cordial

- Supplies specific, positive, informative and corrective error messages
- Prompts user when system is ready for input
- Requests clarification if unable to interpret user's command/response
- Immediately acknowledges receipt of user's command/request
- Keeps user informed of system's processing status or activities
- Provides "point of need" user assistance when requested
- Assumes directive control/initiative only when user requests or permits
- Provides for user's assumption of control any time during the dialogue

III. Protective and Forgiving

- Compensates for common linguistic errors by user
- Accepts common synonyms for command or parameter words
- Requests user confirmation before executing costly, dangerous, or irreversible operations
- Provides user ability to back up and recover after committing an error in logic or the use of a functional capability
- Retains point of progress when user errs or requests specific help
- Enables fast, convenient correction of input errors prior to execution
- Provides standard prompts as reminders of routine procedures
- Prohibits user damaging system or accidentally destroying search progress

IV. Reliable and Responsive

- Scheduled system availability not less than institution's user access periods
- System downtime not more than 2% of scheduled availability
- System actions are regular, predictable, and non-catastrophic
- User's work progress saved if system crashes
- System response time maintained within a range of a few seconds
- System response time for similar transactions remains uniform over time

V. Adaptive and Flexible

- Provides user selection of default values for parameters, display formats, etc.
- Provides user selection of dialogue mode and system message length
- Provides user capability to select content, format, and sorting sequence of display/print
- Provides a modular command language with developmental syntax and "add on" functional extensions to accommodate changing user ability levels
- Monitors user's proficiency with commands and functions, and supplies prompts when need for assistance is detected
- Adapts dynamically with time to accommodate user's changing need for guidance with search strategy/formulation, and provides guidance when appropriate

(Source: Hildreth, 1982b, "Figure 6. Guideline Clusters for the User-Oriented Interface," p. 59)

Figure 38. Guidelines for OPAC Evaluation and Improvement

Hildreth (1982b, p. 59) offers "Guideline Clusters for the User-Oriented Interface" (Figure 38), which he constructed from a comprehensive review of the literature on the design of a user-oriented interface. As new OPACs are developed and existing ones are enhanced, system designers must evaluate them in view of Hildreth's guidelines:

I. Easy to Use
II. Friendly and Cordial
III. Protective and Forgiving
IV. Reliable and Responsive
V. Adaptive and Flexible.

An evaluation of user interfaces could employ the method pioneered by Back (1976) to collect online searchers' spoken thoughts about their online searches, as well as machine transaction logs of searches and searchers' afterthoughts about their searches. Fidel (1981) recently used part of this method to collect online searchers' spoken thoughts. Suggested evaluations could include searchers with varying degrees of online experience and/or training, comparisons of features embedded in different interfaces, different search objectives, etc. The varieties of evaluation exercises are virtually endless; such studies would serve not only to evaluate OPAC designs but to discover the user's model of the information retrieval process which, according to Hildreth (1982b, p. 41), is the crux of the user interface design: "The design of the user interface must be grounded in the user's model of the information retrieval process."

Our knowledge of traditional library catalog searching cannot be relied upon when designing the user interface of OPACs because such an interface was never present in library catalogs. Even though the user interface adds a new dimension to OPACs, system designers cannot ignore Hildreth's guidelines lest they create systems that bear none of the characteristics of a user-oriented interface. For the user-oriented interface is the crucial communications component of the online public access catalog through which the searcher unlocks the system's power.

Figure 39 depicts the necessary components of OPACs with as much or more power than the traditional library catalog. An online public access catalog in which the searcher's model of the information retrieval process of traditional catalog searching is implemented (i.e., one which includes all "required" features disclosed in this report) may have the same potential as the traditional subject catalog, but the inclusion of a user-oriented interface will make that OPAC as powerful as the library catalog. In the absence of a user-oriented interface, such an OPAC supports less than the present search tactics of subject catalog searchers because of the inherent user-oriented nature (e.g., ease of use, familiarity, and unimposing nature) of the library catalog. Likewise, an OPAC, enhanced with features never before available to library catalog searchers (e.g., title and subject keyword access and keyword-in-context vocabulary displays) may have more potential than traditional library catalogs, but the inclusion of a user-oriented interface actually imbues that OPAC with more power than the traditional library catalog. The user-oriented interface is a necessary component of online public access catalogs regardless of the type or number or sophistication of search features or whether the OPAC has as much or more potential than the traditional library catalog.

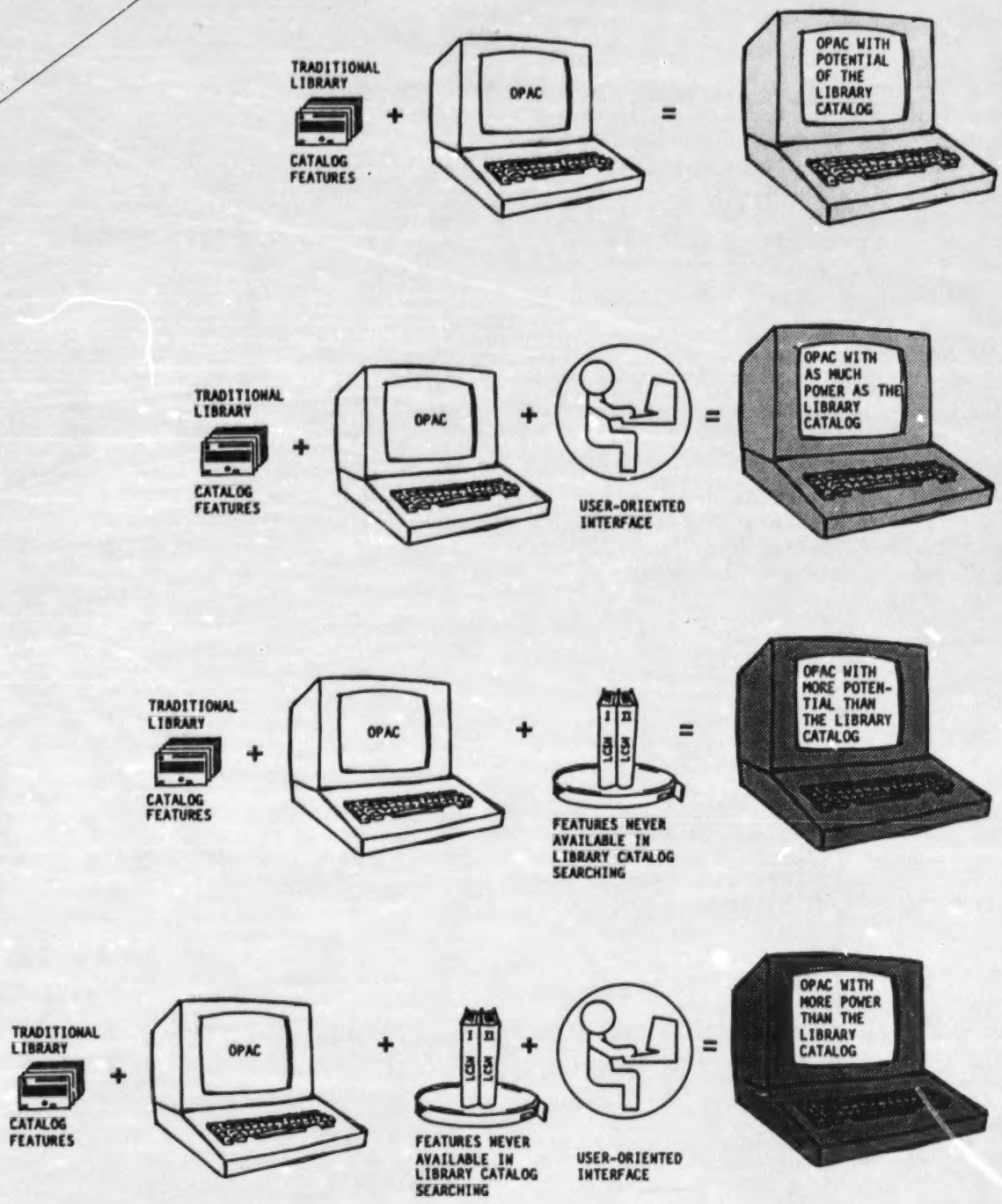


Figure 39. Necessary Components of OPACs with as Much or More Power than Traditional Library Catalogs

REFERENCES

1. Anon. Subject guide to the collection of the Public Library of Columbus and Franklin County. Columbus, OH: Public Library of Columbus and Franklin County; 1978.

2. Atherton, Pauline. Catalog user's access from the researcher's point of view: past and present research which could affect library catalog design. Gapen, D. Kaye; Juergens, Bonnie, eds. Closing the catalog. Proceedings of the 1978 and 1979 Library and Information Technology Association institutes; 1978 November 28-30, 1979 February 14-16; New Orleans, San Francisco. Phoenix, AZ: Oryx Press; 1980: 105-122.

3. Atherton, Pauline; Markey, Karen. Online searching of ERIC: impact of free text or controlled vocabulary searching on the design of the ERIC data base. Pt. 4: the redesign of the ERIC data base for online searching. Syracuse, NY: ERIC Clearinghouse on Information Resources; 1979. ED 180475.

4. Back, Harry Bernard. The design and evaluation of an interactive reference retrieval system for the management sciences. Pittsburgh: Carnegie-Mellon University, Graduate School of Industrial Administration; 1976. 2v. Dissertation.

5. Byers, Dorothy Furber. Effects of cognitive style and information load on process in a decision-making task. Syracuse, NY: Syracuse University, School of Information Studies; 1976. Dissertation.

6. Christ, John M. Concepts and subject headings: their relation in information retrieval and library science. Metuchen, NJ: Scarecrow Press; 1972. 174 p.972.

7. Coblenz, Kathie; Tagliacozzo, Renata. Matching query terms to catalog subject headings. Kochen, Manfred, ed. Integrative mechanisms in literature growth: final report. Ann Arbor, MI: University of Michigan, Mental Health Research Institute; 1970 January: IV:111-126.

8. Cochrane, Pauline A. Subject access in the online catalog. Research Libraries in OCLC: A quarterly. 5:1-7; 1982a January.

9. Cochrane, Pauline A. "Friendly" catalog forgives user errors. American Libraries. 13(5):303-306; 1982b May.

10. Cochrane, Pauline A. Classification as a user's tool in online public access catalogs. Universal classification. Proceedings of the 4th international study conference on classification research; 1982 June 28-July 2; Augsburg. Frankfurt: Indeks Verlag; 1982c: 260-267.

11. Cochrane, Pauline A. Considerations in the design of subject displays for the online public access catalog. Information interaction. Proceedings of the 45th ASIS annual meeting; 1982 October 17-21; Columbus, OH. White Plains, NY: Knowledge Industry Publications; 1982d: 19:359.

REFERENCES

12. Dewey, Melvil. Dewey decimal classification and relative index. 19th ed. Albany, NY: Forest Press; 1979.

13. Doszkocs, Tamas E. Associative indexing, classification, and searching. The information community: an alliance for progress. Proceedings of the 44th ASIS annual meeting; 1981 October 25-30; Washington, DC. White Plains, NY: Knowledge Industry Publications; 1981: 18:329.

14. Fidel, Raya. Operationalist and conceptualist searchers: a case-study-based pattern model of online searching styles. Lunin, Lois F.; Henderson, Madeline; Wooster, Harold, eds. The information community: an alliance for progress. Proceedings of the 44th ASIS Annual Meeting; 1981 October 25-30; Washington, DC. White Plains, NY: Knowledge Industry Publications; 1981: 18:68-70.

15. Freeman, Robert R; Atherton, Pauline. File organization and search strategy using the Universal Decimal Classification in mechanical retrieval systems. Samuelson, Kjell, ed. Mechanized information storage, retrieval, and dissemination. Proceedings of the FID/IFIP joint conference; 1967 June 14-17; Rome. Amsterdam: North-Holland Publishing; 1968: 122-152.

16. Garfield, Eugene. Online and print information services are not always equivalent for all users. Database 3: 4-6; 1980 September.

17. Hildreth, Charles R. The concept and mechanics of browsing in an online library catalog. Proceedings of the 3rd national online meeting; 1982 March 30-April 1; New York. Medford, NJ: Learned Information, Inc.; 1982a: 181-196.

18. Hildreth, Charles R. Online public access catalogs: the user interface. Dublin, OH: OCLC; 1982b. 263 p.

19. Hildreth, Charles R. Online browsing support capabilities. Information interaction. Proceedings of the 45th ASIS annual meeting; 1982 October 17-21; Columbus, OH. White Plains, NY: Knowledge Industry Publications; 1982c: 19:127-132.

20. Horowitz, Gary L.; Bleich, Howard L. PaperChase: a computer program to search the medical literature. New England Journal of Medicine. 305(16):924-930; 1981 October 15.

21. Kaske, Neal K.; Sanders, Nancy P. Online subject access: the human side of the problem. RQ. 20(1):52-58; 1980a Fall.

22. Kaske, Neal K.; Sanders, Nancy P. Evaluating the effectiveness of subject access: the view of the library patron. Benenfeld, Alan R.; Kazlauskas, Edward John, eds. Communicating information. Proceedings of the 43rd ASIS annual meeting; 1980 October 5-10; Anaheim, CA. White Plains, NY: Knowledge Industry Publications; 1980b: 17:323-325.

REFERENCES

23. Kugel, Peter. Dirty Boole: computer-based document retrieval systems. Journal of the American Society for Information Science. 22: 293-294; 1971 July-August.

24. Lancaster, Frederick Wilfrid. Information retrieval systems: characteristics, testing, and evaluation. New York: Wiley; 1968. 222 p.

25. Markey, Karen. Analytical review of catalog use studies. Columbus, OH: OCLC; 1980; OCLC research report OCLC/OPR/RR-80/2. Available from: ERIC; ED 186041.

26. Mischo, William H. A subject retrieval function for the online union catalog. Columbus, OH: OCLC; 1981; OCLC technical report OCLC/DD/TR-81/4. Available from: ERIC; ED 212263.

27. Penniman, William David. Rhythms of dialogue in human-computer conversation. Columbus, OH: The Ohio State University, Department of Communication; 1975. 281 p. Dissertation.

28. Summit, Roger K. Popular illusions relating to the costs of online services. Paper read at the 99th annual meeting of the American Library Association; 1980 June 28-July 4; New York. Quoted in American Libraries. 425; 1980 July-August.

29. Tagliacozzo, Renata. Types of catalog searches and their relationship to some characteristics of the users. Kochen, Manfred, ed. Integrative mechanisms in literature growth: final report. Ann Arbor, MI: University of Michigan, Mental Health Research Institute; 1970 January: IV:27-39.

APPENDIX A

Agenda: Training For Protocol Analysis Interviews
October 1979
Dorothy F. Byers, OCLC Consultant

8:30 - 9:00 AM Some basics of interviewing
9:00 - 10:30 AM Practicing basic techniques
10:30 - 11:00 AM Break
11:00 - 12:00 Look at your videotapes and/or listen to tapes
12:30 - 1:30 PM Lunch
1:30 - 2:00 PM Problems peculiar to this task
2:00 - 3:30 PM Practice
3:30 - 4:00 PM Break
4:00 - 5:00 PM Analyze tapes from practice and improve

8:30 - 9:00 BASICS OF INTERVIEWING

I. Two major purposes in interviewing

A. Obtaining quality of participation
B. Obtaining quality of information

II. Achieving those purposes

A. Fostering participation

1. Importance of interviewer's role
2. Reward
3. Nonapologetic attitude (do not ask for cooperation except with university students)
4. Good statement of purpose--what you are doing, what you want to do
5. Appeal to respondent
6. What's involved (especially time)
7. Sanction (OCLC sign)
8. Handling resistance

B. Getting clear, relevant, complete information

1. Be clear about what you want
2. Use open (indirect) and closed (yes/no) questions appropriately
3. Base questions on respondent's antecedents
4. Probe

9:00 - 10:30 PRACTICE (sample, practice, tape, review, practice)

I. Introduction

A. Hello. I'm working on a survey to find out how people use the card catalog so that it can be improved. I'd like to have you think out loud while you use the catalog. I'm going to use a tape recorder so I don't have to take alot of notes, is that O.K.? Thank you, go ahead.

B. Don't be apologetic: Hi. I am part of a research investigation to examine the way people do a subject search in the card catalog. Do you suppose I could ask you to take part in this project? It

APPENDIX A

involves having you talk out loud while you search the catalog. Is it O.K. with you if I tape what you say?

II. Handling resistance

A. "I'm too busy now"
B. "I don't know anything about that"
C. "I don't want to be on tape"
D. "Why are you taking notes"
E. "What do you think?"

III. Fostering participation

A. Role of interviewer
B. Reward - better service, cognitive or intellectual kicks
C. What's involved - not much more time than plain search
D. Sanction - sign by library announcing presence of project

IV. Probing

A. Repeat question or statement
B. Pause
C. Repeat Respondent's statement
D. Neutral comments or questions:

1. "Anything else?"
2. "I see"
3. "You're looking at ...?"
4. "What part of the card are you looking at?"
5. "Why did you flip past that one?"

E. Encourage
F. Giggles and interruptions

V. Post-interview conversation -- ego-boosting comments

A. "That was an interesting search"
B. "You were a terrific respondent"
C. "Did you think that was hard?"
D. "Thanks"
E. Get demographic data from patron before he/she leaves

10:30 - 11:00 BREAK

11:00 - 12:30 TAPES AND VIDEO TAPES

I. Discuss any interviewing problems - e.g., of good and bad
II. Problems which will affect interviewing process
III. Possible DV's

12:30 - 1:30 LUNCH

APPENDIX A

1:30 - 2:00 PROBLEMS WITH THIS TASK

I. Using the tape recorder as R (respondent) moves
II. Recording what R does physically
III. Getting R to speak clearly
IV. Getting R to speak
V. Preventing R from looking to interviewer (for help, ideas, etc.)
VI. Conducting open schedule with closed procedure

2:00 - 3:30 PRACTICE

I. Practice using tape recorder
II. Practice note-taking and juggling with recorder
III. Practice prompting R
   A. "Talk louder"
   B. "You're not talking"
   C. "What are you thinking?"
   D. "What are you doing now?"
   E. "What did you say?"

IV. Practice and record whole sequence from introduction to post-interview

3:30 - 4:00 BREAK

4:00 - 5:00 REVIEW TAPES FROM IV ABOVE AND PRACTICE

APPENDIX B

Protocol Analysis Interview Questionnaire

PATRON INTERVIEWS TAPE # ___ SIDE # ___

LOCATION _____ INTERVIEW # _____

_____

__ Public __ School __ Special __ College __ University

Physical movements noted _____

_____

TYPE OF CATALOG __ Dictionary __ Divided A/T & S

__ Divided A, T, & S __ Other _____

Items/terms brought to the catalog (notes, names, titles)

_____

DATE 1980 _____

INTERVIEWER _____

(Questions to be used at the end of a patron interview.)

With recorder on, have patron state purpose of this catalog use/catalog search if you do not yet know it. If you do, confirm your knowledge by asking the following question:

"If I understand right, you have been using the catalog to find materials on ... such & such ..., is that correct?" __ Yes __ No

(Note patron's purpose of this catalog use.)

_____

_____

_____

"You came to the library today in order to?"

(Note purpose of patron's current library visit.)

_____

_____

_____

___ Yes __ No

APPENDIX B

State to interviewee: ("To complete my work I need to ask you a few short questions.")

1. When were you last in this library?

   __ never before __ days ago __ weeks ago __ months ago __ years ago

2. How often do you go to libraries?

   Every so many __ days __ weeks __ months __ years

3. What libraries do you use in addition to this one?

   __ Public __ College __ University __ School __ None

   __ Special Type _____ __ Jr. College

   __ Church __ Other _____

4. Have you ever had a formal class or lecture on how to use a library catalog?

   __ Yes __ No

   Do you remember when and where that took place?

   _____ Grade School __

   _____ High School __

   _____ Jr. College __

   _____ College __

   _____ Public Library __

   _____ Other __

State to interviewee: ("From this card please give me your age group and years of schooling.")

5. Sex __ Male __ Female

6. Age Group ____

7. Schooling Group ____

   Degrees _____ Field _____

   Degrees _____ Field _____

   Degrees _____ Field _____

   Degrees _____ Field _____

THANKS FOR YOUR COOPERATION IN THIS STUDY!

APPENDIX C1. Expressed Topic and Corresponding Initial Access Point as Partial Representation of the Former (i.e., 1A)

Expressed Topic	No. of Terms	Initial Access Point (1A)	No. of Terms
Psychological Aspects of Rape	3	Rape	1
Modern Finance	2	Finance	1
Social Work	2	Social Services	2
Poussin, Nicholas	2	Poussin	1
Black Art Museums in the U.S.	4	Afro American	2
Misinterpretation of Artwork by Critics	3	Critics	1
Effect of Politics on Byzantine & Early Christian Architecture	6	Architecture, Byzantine	2
Rothko & Pollock	2	Rothko	1
Tennessee Williams Play	3	Williams, Tennessee	2
Comparison of U.S. & U.S.S.R. Defense Capability	5	Defense	1
Maps of Ohio	2	Maps of U.S.	2
Teenage Alcoholism & Alcohol Rehabilitation Programs	5	Alcoholism	1
Discrimination in Employment Tests	3	Discrimination	1
Financial Statement Analysis	3	Financial Statements	2
Mountain Men	2	Mountain	1
Costumes in Cycle Plays	3	Theater--Costuming	2
Selecting a Moving Van Company	4	Moving	1
Watercolor Techniques	2	Watercolors	1
Hot Air Balloons	3	Balloons	1
Criticism of Susan Brownmiller's "Against Our Will"	4	Brownmiller, Susan	2
Reinforcement to Stop Smoking	3	Smoking	1
Descriptions of Occupations	2	Occupations	1
Poems of Elizabeth Barrett Browning	4	Browning	1
Satanic Element in "The Picture of Dorian Gray"	3	The Picture of Dorian Gray	1
Three Approaches to Schizophrenia	3	Schizophrenia	1
Trance Dances of the Holy Ghost People	5	Trances	1
Careers in Banking	2	Bank	1
Commerce & Geography of Poland	3	Poland	1
H.G. Well's Theory of Social Realism	6	Wells, H.G.	3
Artistic Photography	2	Photography	1
Commentary on Plato's "Republic"	3	Plato	1

APPENDIX C1

Expressed Topic	No. of Terms	Initial Access Point (1A)	No. of Terms
Political Advertising on Television	3	Political Advertising	2
World War II Tanks	4	World War, 1939-1945	3
Effects of Hunting on Animal Ecology	4	Animals	1
History of Soccer	2	Soccer	1
Life After Death	3	Life	1
Beginning Chemistry	2	Chemical Bonds	2
Africa, Natural Resources and Their Effects	4	Africa	1
Basic Engineering Book	3	Engineering	1
Russian Sports	2	Sports	1
Infant and Maternal Health in China, India, & the U.S.	6	China	1
Hair Disorders	2	Hair	1
Self-Concept of Women after Divorce	4	Divorce	1
Pediatric Optometry	2	Optometry	1
Establishing Ambulatory Care Programs	4	Ambulatory Medical Care	3
Morrocan Design	2	Morroco	1
Binding and Psychological Effects of Separation on Twins	5	Twins	1
Victorian Homes	2	Victorian Architecture	2
Civil War Medicines	3	Civil War	2
F. Scott & Zelda Fitzgerald	4	Fitzgerald, F. Scott	3
Gibson Girls	2	Gibson	1
Extrasensory Perception	2	Extrasensory	1
Mexican Sociology & Social Conditions	4	Mexico	1
Different Types of Smokers and Reasons for Smoking	5	Smoking	1
Guide to Writing Resumes	3	Resumes	1
Kennedy Assassination	2	Kennedy, J.	1
Controlling Management	2	Management	1
Timing Cars for High Performance	4	Automobiles	1
Small Business	2	Business	1
Christmas in Germany	2	Christmas	1
Description of Florida's East Coast	4	Florida - Description and Travel	3

APPENDIX C1

Expressed Topic	No. of Terms	Initial Access Point (1A)	No. of Terms
Roman History	2	Roman	1
Personal Income Tax	3	Income Tax	2
Wedding Photography	2	Photography	1
How Locks Work	3	Locks	1
Rules & Regulations Governing Small Business	5	Small Business	2
Effect of Multinational Corporations on 3rd World Countries	6	Multinational Corporations	2
Canadian Anthem	2	Canada	1
Mimes	1	Mim	1
History of Baseball	2	Baseball	1
Mountain Climbing Ethics	3	Mountaineering	1
Geology of the Adirondacks	2	Adirondacks	1
TOTAL =	223	TOTAL =	97
AVERAGE =	3.1	AVERAGE =	1.35

APPENDIX C2. Expressed Topic and Corresponding Middle or Last Access Point as Partial Representation of the Former (i.e., 1A)

Expressed Topic	No. of Terms	Middle or Last Access Point (1A)	No. of Terms
Arizona	1	Arizona University	2
Tree Forts	2	Forts	1
Financing of Schools	2	School	1
Financing of Schools	2	Finance	1
Life after Death	2	Future Life	2
Symbol for No Left Turn	3	Signs	1
Drag Racing	2	Automobile Racing	2
Chivalry at Arthur's Round Table	4	Arthur	1
Infant and Maternal Health in China, India, and the U.S.	6	Maternal Health Services	3
Pediatric Optometry	2	Pediatric Opthamology	2
Civil War Medicines	3	U.S. History--Civil War	4
Racquetball and Bowling	2	Bowling	1
Racquetball and Bowling	2	Racquetball	1
Photovoltaic Cells	2	Photoelectric Cells	2
Tax Exempt Status of Churches	4	Tax	1
Acid Rain and Military Draft	4	Draft	1
Controlling Management	2	Control	1
Executive and Corporate Stress	2	Stress	1
Introductory Chemistry	2	Chemistry	1
Abbey Theatre	2	Dublin. Abbey Theatre	3
Black Art Museums in the U.S.	4	Black American	2
Black Art Museums in the U.S.	4	Museums	1
Frank Stella and Kenneth Noland	4	Noland, Kenneth	2
Rothko and Pollock	2	Pollock	1
Television Audience Ratings	3	Television	1
Television Audience Ratings	3	Audience	1
Commerce and Geography of Poland	3	Poland--Description and Travel	3
Commerce and Geography of Poland	3	GNP--Poland	2
Political Advertising on Television	3	Advertising, Political	2
Political Advertising on Television	3	Television in Politics	2
Wedding Photography	2	Wedding	1
TOTAL =	85	TOTAL =	50
AVERAGE =	2.74	AVERAGE =	1.61

APPENDIX C3. Number of Terms in Expressed Topics and Middle or Last Access Points--1A: t-Test for Two Related Samples

Null Hypothesis: Mean of the number of terms in expressed topics is equal to the mean of the number of terms in middle or last access points--1A.

Alternate Hypothesis: Mean of the number of terms in expressed topics is greater than the mean of the number of terms in middle or last access points--1A.

t-test for related samples, N=31
.005 level, one-tailed, degrees of freedom = 30
R:$t \geq 2.75$

$\overline{D} = 1.129$

$$\Sigma d^2 = 81 - \frac{35^2}{31} = 41.4839$$

$$S_{\overline{D}} = \sqrt{\frac{41.4839}{31(31-1)}} = .2112$$

$$t = 1.129/.2112 = 5.346^*$$

*Rejects the null hypothesis at .005, 30 degrees of freedom

Thus, the number of terms in expressed topics is significantly greater than the number of terms in middle or last access points--1A.

APPENDIX C4. Expressed Topic and Corresponding Initial Access Point as Different from the Former (i.e., 1C)

Expressed Topic	No. of Terms	Initial Access Point (1C)	No. of Terms
Treatment of Blacks before 24th Amendment	4	Slavery	1
Financing of Schools	2	Federal Aid to Education	3
Sitting Bull	2	Indians of North America	3
Symbol for No Left Turn	4	Automobiles	1
Monocytes	1	Blood	1
Boat People	2	Vietnamese	1
Soil Mechanics and Building	3	Pilings	1
Women, Infant, and Children Agency	4	Poor	1
Breakfast Goodness	2	Nutrition	1
Tax Exempt Status of Churches	4	Religion	1
U.S. Air Force	3	Military Arts	2
Togas	1	Costumes, Ancient or Roman	3
Executive and Corporate Stress	3	Mental Illness	2
Introductory Chemistry	2	Chemical	1
Frank Stella and Kenneth Noland	4	Abstract Expressionism	2
Lawrence Kohlberg's Theory of Moral Education	5	Moral Education	2
German Language	2	Germany	1
Wilmington Oil Fields	3	Los Angeles	2
Financial Ratios	2	Accounting	1
Plains Rifle	2	Russell	1
Player Piano	2	Keyboard Instruments	2
FCC	1	U.S. Government	2
TOTAL =	58	TOTAL =	35
AVERAGE =	2.64	AVERAGE =	1.59

APPENDIX C5. Expressed Topic and Corresponding Middle or Last Access Point as Different from the Former (i.e., 1C)

Expressed Topic	No. of Terms	Middle or Last Access Point	No. of Terms
NATO	1	North American Treaty Organization	4
Financing of Schools	2	Education	1
Life After Death	2	Immortality	1
Whale Bone	2	Fable	1
Symbol for No Left Turn	3	Law	1
American Research Bureau	3	Television	1
Fables	1	Folklore	1
Monocytes	1	Pathology	1
Monocytes	1	Physiology	1
Monocytes	1	Leukocytes	1
Dreams	1	Sleep	1
Munch	1	Munch	1
Munch	1	Expressionism--Art	2
Effects of Snuff	2	Tobacco--Smoking	2
U.S.S.R.	1	Russia	1
Infant and Maternal Health in China, India, & the U.S.	3	Infant Mortality	2
Establishing Ambulatory Care Programs	4	Hospitals--Outpatient Services	3
Women, Infant, and Children Agency	4	Public Health--Research	3
Women, Infant, and Children Agency	4	Nutrition	1
Crystallography	1	Diamonds	1
Crystallography	1	Precious Stones & Gems	3
Crystallography	1	Mineralogy	1
Water Safety	2	Scuba	1
Afghan Skin Problems	3	Dogs	1
Separation Anxiety	2	Child Development	2
Female Impersonators	2	Homosexuality	1
Aycock, Alice	2	Artists, American	2
Aycock, Alice	2	Sculpture	1
Black Art Museums in the U.S.	4	Negro	1
Black Art Museums in the U.S.	4	New York City	3
Black Art Museums in the U.S.	4	Harlem	1

APPENDIX C5

Expressed Topic	No. of Terms	Middle or Last Access Point	No. of Terms
Misinterpretation of Artworks by Critics	3	Criticism	1
The Draft or Military Conscription	3	World War II, 1939-1945	3
Blackjack	1	Cards	1
Blackjack	1	Games--Rules	2
Costumes in Cycle Plays	3	Theater	1
Costumes in Cycle Plays	3	Theater--Moral & Religious Aspects	4
Wilmington Oil Fields	3	Geology	1
Satanic Element in "The Picture of Dorian Gray"	4	Wilde, Oscar	2
Trance Dances of the Holy Ghost People	5	Holy Spirit	2
Trance Dances of the Holy Ghost People	5	Religion, Primitive	2
Television Audience Ratings	3	Oligopoly	1
Television Audience Ratings	3	T.V. Advertising	2
Audi Repair Manual	3	Automobile	1
Lunch Bag - How to Make	3	Sewing	1
Crockpot Cooking	2	Cookery	1
Pinatas	1	Mexico	1
Pinatas	1	Lewis, Oscar	2
Geology of the Adirondacks	2	Natural Resources	2
How to be a Church Historian	3	Historical	1
FCC	1	Media	1
FCC	1	Mass Media	1
TOTAL =	123	TOTAL =	80
AVERAGE =	2.37	AVERAGE =	1.54

APPENDIX C6. Number of Terms in Expressed Topics and Middle or Last Access Points--1C: t-Test for Two Related Samples

Null Hypothesis: Mean number of terms in expressed topics is equal to the mean of the number of terms in middle or last access points--1C.

Alternate Hypothesis: Mean number of terms in expressed topics is greater than the mean of the number of terms in middle or last access points --1C.

t-test for related samples, N=52
.005 level, one-tailed, degrees of freedom = 51
R:t $\geq$ 2.704

$\overline{D}$ = .7885

$$\Sigma d^2 = 118 - \frac{41^2}{52} = 85.6731$$

$$S_{\overline{D}} = \sqrt{\frac{85.6731}{52(52-1)}} = .1797$$

$$t = .7885/.1797 = 4.387^*$$

*Rejects the null hypothesis at .005, 51 degrees of freedom.

Thus, the number of terms in expressed topics is significantly greater than the number of terms in middle or last access points--1C.

APPENDIX C7. Expressed Topic and Corresponding Initial Access Point Same as the Former (i.e., 1B)

Expressed Topic & Initial Access Point (1B)	No. of Terms
Marijuana	1
Prostitution	1
Faith Healing	2
Milky Way	2
Franklin, Benjamin	2
Chicago Fire	2
Henry, Patrick	2
Fables	1
Dogs	1
Pharmacy	1
Philip II of Macedonia	3
Prohibition	1
Nutrition	1
Murder	1
Prisons	1
Information Theory	2
Electromagnetic Theory	2
Jackson, Andrew	2
Egypt	1
Canada	1
Iran	1
Russia	1
Germany	1
Mythology	1
Crystallography	1
Abortion	1
Italy	1
Advertising	1

APPENDIX C7

Expressed Topic & Initial Access Point (1B)	No. of Terms
Treatment of Alcoholism	2
Animation	1
Abbey Theatre	2
Mondrian	1
Beuys, Joseph	2
Biofeedback	1
Genghis Khan	2
Chamber of Commerce	2
Dog Breeding	2
Hypertension	1
Repairs	1
Kenton, Simon	2
Calligraphy	1
Surveys	1
Real Estate	2
Renoir, Jean	2
Prenatal Care	2
Backpacking	1
Greece	1
Airline Attendant	2
Cacti and Succulent Plants	3
Spinning Wheels	2
Etiquette	1
Anorexia Nervosa	2
Bargello	1
Motorcycles	1
	TOTAL = 79
	AVERAGE = 1.46

INDEX

Access points, 3-4, 21, 28-29, 42-60, 79-80, 124-134
  and assistance, 4, 43, 47, 75, 80
  and participating libraries, 4, 59, 80
  and refinements, 5, 69, 71-73
  and scans, 62, 66
  in bookshelf browsing searches, 3, 7, 38-39, 62, 66, 79
  by constituent elements, 21-23, 31, 42-47, 59
    different from expressed topic, 4, 5, 21, 42, 49, 52-54, 69, 72, 79, 84, 92
    exact representation of expressed topic, 4, 21, 42-43, 52, 55-58, 72, 79
    partial representation of expressed topic, 4-5, 21, 42-43, 48-51, 69, 72, 79, 84, 92
  definition of 3, 19, 42-43
  examples of, 21, 42-43, 48-49, 52
  initial, 4, 45-60, 69, 71-73, 79-80, 83
  last, 43, 45, 47
  middle, 43, 45, 47
  number of, 3-4, 8, 31, 38-39, 43-47, 62, 66, 79, 100
  number of terms in, 4, 7, 48-56, 59-60, 69, 79
  in process model, 83, 85, 89, 94, 102, 104
  protocols featuring, 51, 54, 57-58 64-65, 101
Ad hoc process models. See Process models
Aid from librarian. See Librarian assistance
Assistance. See Librarian assistance
Atherton, Pauline A. See Cochrane, Pauline A.
AUDACIOUS Project (UDC), 111, 117

Back, Harry Bernard, 16, 114, 116
Bleich, Howard L., 109, 117
Bookshelf browsing, 3, 7, 29, 35-38, 62, 77, 79, 80, 84-87
  characteristics of, 7, 29, 38, 79, 84
  process model of, 7, 85
  protocol featuring, 35, 41
  required OPAC features for, 7, 84, 86
  suggested OPAC features for, 7, 35-38, 84, 86-87
Boolean operations, 5, 9, 69, 72, 75, 80, 82, 87, 91, 95-99
Brownrigg, Edwin B., 75
Browsing, 5, 68, 70, 72, 80
Byers, Dorothy F., 16, 116

Catalog, type of, 26, 99, 122-123
Christ, John M., 16, 116
CITE (OPAC at National Library of Medicine), 10, 99, 111
Claremont Colleges' OPAC (TLS), 5, 9, 69, 80, 91, 95
Classified lists or logical outlines, 7, 8, 37, 109-111
Coblenz, Kathie, 48, 116
Cochrane, Pauline A., 10, 36, 92, 105, 109-112, 116-117
Constituent elements, 2, 6, 21-23, 79, 82
  and bookshelf browsing, 7, 38-39, 84
  and participating libraries, 38, 40
  number of, 3, 28-29, 31-32, 34, 38-39, 79, 92
  protocols featuring, 24, 33
Cross references, 8, 80, 100, 105-106, 111

Dewey Decimal Classification, 8, 36-38, 42, 79, 110-111, 117
DIALOG retrieval system, 9, 92
Displays, 12, 62-63, 77, 80
  brief bibliographic records, 11, 62
  full bibliographic records, 6, 77
  intermediate bibliographic records, 11, 62-63, 80
Doszkocs, Tamas E., 99, 117

Exit, 20, 23, 29, 32, 77
  in process model, 83, 85, 89, 94, 102, 104
Expressed topics, 4, 42, 50, 53, 79, 111, 124-134
  examples of, 21, 48-49, 52
  in process model, 83, 85, 89, 94, 102, 104

Features for OPACs. See Required OPAC features, Suggested OPAC features
Feedback, 10, 109-111
Fidel, Raya, 114, 117
Flags, 6, 11, 20, 22, 29, 32, 77
Freeman, Robert R., 111, 117

Garfield, Eugene, 75, 117

Hildreth, Charles R., 5, 13, 72, 75, 91-92, 99, 100, 105, 112-114, 117
Hits. See Postings
Horowitz, Gary L., 109, 117

Kaske, Neal K., 14, 117
Kugel, Peter, 75, 118

Lancaster, F. W., 5, 72, 118
LCS (Library Control System, OPAC of The Ohio State University), 84, 86, 100, 105
Librarian assistance, 4, 6, 20, 22, 29, 32, 43, 47, 75-77, 80, 100, 103-105
  in process model, 83, 104
  protocol featuring, 103
Library of Congress Subject Headings, 5, 8, 60-61, 100, 105-106
Limits, 6, 9, 75, 80, 91, 95, 111

Maggie's Place (OPAC of the Pike's Peak Library District), 87, 90
Mankato State University's OPAC (MSUS/PALS), 9, 91, 95-97, 99
Markey, Karen, 3, 14, 109, 118
MELVYL (OPAC of University of California), 10, 91-92, 95, 99, 100, 105, 107-108
Methodology, 1, 16-24, 119-123
Mischo, William H., 92, 99, 109, 118

Objectives of the Project, 1, 14-15
OPACs. See CITE, Claremont Colleges' OPAC, LCS, Maggie's Place, Mankato State University's OPAC, MELVYL, PaperChase, SULIRS
ORBIT retrieval system, 5, 9, 69, 80, 91, 95

INDEX

PaperChase (OPAC of Beth Israel Hospital), 10, 109
Participating libraries, 2, 16, 19, 25-28, 79, 80, 122-123
  and bookshelf browsing, 7, 38, 40, 79
  and constituent elements, 38, 40
  and number of terms in access points, 4, 56, 59-60
Penniman, W. David, 1, 19, 118
Perseverance, 3, 4, 43-44, 79
Postings, 11, 77, 84, 99, 105
Power, 1, 12-15, 82, 114-115
Process model, 1, 14-15, 17, 82-85, 87, 89, 92, 94, 102, 104
  of bookshelf browsing, 7, 84-86
  composite, 6, 82-83
  of failure, 100-102
  with librarian assistance, 100-111
  with refinements, 9, 87-99
  requiring vocabulary assistance, 7-8, 100-111
Prompts, suggestive, 9, 12, 42, 62-63, 99, 105
Protocol analysis, 1, 14, 16-18, 25, 119-121
Protocols, 17-19, 23, 79
  by topic, Airships, 68
    Audi repair manual, 88
    Canada's anthem, 51
    Christmas in Germany, 33
    FCC, 101
    Financial statement analysis, 24
    How good a breakfast is, 54
    Information theory, 35
    Life after death, 103
    Mondrian, 58
    Motorcycles, 57
    Multinational corporations, 93
    Prenatal care, 70
    Real estate, 41
    Small business, 74
    Trance dances of the Holy Ghost People, 64

Records, 6, 29, 77-80, 84
  and bookshelf browsing, 7, 80, 84
  by constituent element, 23, 32, 78
    author, 77-78
    call number or call number areas, 6, 11, 29, 77-80, 84
    full bibliographic information, 6, 77-78, 80
    title, 77-78
  definition of, 20, 23
  in process model, 83, 85, 94, 104
  protocol featuring, 101
Refinements,
  a priori, 5-6, 19, 28, 62, 67-75, 80, 87-92
    and access points, 5, 71, 80
    as makeshift Boolean operations, 5, 9, 68-70, 72, 75
    by constituent element, 21, 31
    definition of, 5, 19, 62, 87
    number of, 31, 62-63, 69, 71, 75, 92
    in process model, 83, 89, 94
    protocol featuring, 68, 70, 74, 88
    for relevance, 62, 69
    required OPAC features for, 9, 91
    suggested OPAC features for, 9-10, 91-93
  interactive, 5-6, 19, 29, 62, 67-75, 80, 92-99
    and access points, 5, 80
    as limits, 6, 9, 75, 80, 91, 95
    as makeshift Boolean operations, 5, 9, 69-70, 72, 75
    by constituent element, 22, 32
    definition of, 5, 19, 62, 87
    number of, 32, 62-63, 72-73, 75
    in process model, 83, 94, 104
    protocol featuring, 70, 74, 93
    required OPAC features for, 9, 95-99
    suggested OPAC features for, 9-10, 99
Required OPAC features, 1, 6-15, 17, 82, 84, 86, 91, 95-100, 114-115
  alphabetical subject heading displays, 7, 12, 77, 80, 84
  Boolean operations, 9, 91, 95-99
  browse, backward, 11, 38, 42, 77
    forward, 11, 38, 42, 77
  class number areas, 7-8, 84, 105, 109-111
  cross references, 8, 77, 80, 100, 105, 111
  displays of bibliographic records, 11-12, 62-63, 80
    tracings in, 10-11, 62-63, 80
  "history" or "display sets" command, 11, 77
  limits, 9, 91, 95, 111
  postings, 11, 77, 84, 86, 105
  proximity, of Library of Congress Subject Headings, 8, 105
    of reference librarian, 8, 105
  suggestive prompts in highly posted searches, 9, 99

Sanders, Nancy P., 14, 117
Scans, 5, 12, 29, 38, 60-66, 84
  and bookshelf browsing, 7, 62, 79
  by constituent element, 22, 31, 61
    author, 60-61
    call number, 61-62
    Library of Congress Subject Headings, 5, 60-61, 100, 105-106
    "see also" reference, 60-61
    "see" reference, 60-61, 100-101
    subject headings, 60-61
    titles, 5, 60-61
    tracings, 5, 10, 60-61
  definition of, 5, 19, 60
  number of, 60-62, 66
  in process model, 83, 85, 89, 94, 102, 104
  protocol featuring, 64-65
  of subject-rich information, 5, 11, 60, 62
Search activity codes, 19-23, 28-29, 31-32
  definition of, 1-2, 19-23
  number of, 3, 31-32
    average, 3, 28-29
Searchers, 17-18, 25, 27-28, 79, 122-123
  ages of, 2, 18, 25, 27, 79
  catalog instruction of, 3, 18, 28, 30, 79
  frequency of library use of, 2, 18, 27-28
  purpose of, 3, 18, 28, 30
  years of schooling, 18, 27, 79
Searches. See Protocols
Subject Guide to the Collection of the Public Library of Columbus and Franklin County, 7, 35-36, 38, 42, 79, 86-87
Subject heading displays,
  alphabetical, 7, 12, 77, 80, 84
  in bibliographic records, 10, 62-63, 80
  keyword-in-context, 8, 10, 77, 80, 92, 99, 105, 107-108, 111
  related term, 8, 105-106
  rotated, 9, 92, 99, 109
Subjects. See Searchers

INDEX

Suggested OPAC features, 6-15, 17, 82, 84, 86-87, 91-92, 99, 105, 111, 114-115
  classified lists or logical outlines, 7-8, 37, 109-111
  feedback about relevant output, 10, 109
  feedback about relevant subject headings, 10, 109-111
  keyword-in-context subject heading displays, 8, 10, 77, 80, 92, 99, 105, 107-108, 111
  limit options, 111
  related term displays, 8, 105-106
  rotation of subject headings, 9, 92, 99, 109
  Subject Guide, 7, 35-36, 38, 42, 79, 86-87
SULIRS (OPAC of Syracuse University), 9, 91, 95, 98-99
Summit, Roger K., 75, 118

Tagliacozzo, Renata, 36, 38, 48, 118

User-oriented interface, 12-13, 15, 82, 111-115